IRAS VERSUS POTENT DENSITY FIELDS ON LARGE SCALES: BIASING AND OMEGA

Yair Sigad [1], Amiram Eldar [1], Avishai Dekel [1,2,3],
Michael A. Strauss [4,5], Amos Yahil [6]

ABSTRACT

The galaxy density field as extracted from the *IRAS* 1.2 Jy redshift survey is compared to the mass density field as reconstructed by the POTENT method from the Mark III catalog of peculiar velocities. The reconstruction is done with Gaussian smoothing of radius $12\,h^{-1}$Mpc, and the comparison is carried out within volumes of effective radii $31-46\,h^{-1}$Mpc, containing $\approx 10-26$ independent samples. Random and systematic errors are estimated from multiple realizations of mock catalogs drawn from a simulation that mimics the observed density field in the local universe. The relationship between the two density fields is found to be consistent with gravitational instability theory in the mildly nonlinear regime and a linear biasing relation between galaxies and mass. We measure $\beta_I \equiv \Omega^{0.6}/b_I = 0.89 \pm 0.12$ within a volume of effective radius 40 h^{-1}Mpc, where b_I is the *IRAS* galaxy biasing parameter at $12\,h^{-1}$Mpc. This result is only weakly dependent on the comparison volume, suggesting that cosmic scatter is no greater than ± 0.1. These data are thus consistent with $\Omega = 1$ and $b_I \simeq 1$. If $b_I > 0.75$, as theoretical models of biasing indicate, then $\Omega > 0.33$ at 95% confidence. A comparison with other estimates of β_I suggests scale-dependence in the biasing relation for *IRAS* galaxies.

Subject headings: cosmology: theory — cosmology: observation — dark matter — galaxies: distances and redshifts — galaxies: formation — galaxies: clustering — large-scale structure of universe

[1]Racah Institute of Physics, The Hebrew University, Jerusalem 91904, Israel

[2]Center for Particle Astrophysics, University of California, Berkeley, CA 94720

[3]Physics/Lick Observatory, University of California, Santa Cruz, CA 95064

[4]Princeton University Observatory, Princeton University, Princeton, NJ 08544

[5]Alfred P. Sloan Foundation Fellow

[6]Dept. of Physics & Astronomy, State University of New York, Stony Brook, NY 11794-3800

1. INTRODUCTION

A comparison of the galaxy density field derived from a redshift survey with the mass-density fluctuation field inferred from galaxy peculiar velocity data, allows one to test both gravitational instability theory (GI) and models for the relation between galaxies and mass. If the data are consistent with the assumed model, one can then estimate the value of the cosmological density parameter Ω. This can be illustrated, for simplicity, by the linear approximation to GI, for which the relation between the mass density fluctuation field $\delta(\boldsymbol{x}) \equiv [\rho(\boldsymbol{x}) - \bar{\rho}]/\bar{\rho}$ and the peculiar velocity field $\boldsymbol{v}(\boldsymbol{x})$ is

$$\boldsymbol{\nabla} \cdot \boldsymbol{v} = -f(\Omega)\,\delta \ , \quad f(\Omega) \approx \Omega^{0.6} \ , \quad |\delta| \ll 1 \ , \tag{1}$$

with distances measured in $\mathrm{km\,s^{-1}}$ (*i.e.*, the Hubble constant is set to unity). Observations of galaxy peculiar velocities allow us to measure the quantity on the left-hand side, while galaxy redshift surveys provide a measure of the *galaxy* density fluctuation field, $\delta_g(\boldsymbol{x})$. The latter need not be identical to $\delta(\boldsymbol{x})$ (*cf.*, Bardeen *et al.* 1986; Dekel & Rees 1987); we adopt here the simplest toy model relating the two fields, linear biasing,

$$\delta_g(\boldsymbol{x}) = b\,\delta(\boldsymbol{x}) \ , \tag{2}$$

where b refers to the galaxies in a specific redshift survey and at a fixed smoothing length. With this model for biasing, equation (1) can be rewritten as a relation between the observable quantities,

$$\nabla \cdot \boldsymbol{v} = -\beta\,\delta_g \ , \quad \beta \equiv f(\Omega)/b \ . \tag{3}$$

Thus, in the context of linear GI and linear biasing, the comparison of peculiar velocities and the galaxy distribution enables a measurement of β. However, β provides only an indirect estimate of Ω, because it is a degenerate combination of Ω and b.

There is quite an extensive literature on the comparison of peculiar velocity and redshift survey data to measure β (for reviews see Dekel 1994; Strauss & Willick 1995; Strauss 1997b; Dekel 1997ab; Dekel, Burstein & White 1997 for a general review of Ω measurements). The comparison is not straightforward: the peculiar velocity data are sparse, inhomogeneously distributed, limited to the radial component, and quite noisy, while the redshift data need to be corrected to real space and may trace the mass distribution in a non-trivial way. These difficulties give rise to a variety of statistical biases which depend on the details of the analysis carried out.

There are two approaches to this problem, depending on whether the quantities that are actually compared are velocities or densities. Integrating both sides of equation (3) yields a predicted velocity field given measurements of the galaxy density field (equation 18 below). Comparison of these predictions to observed radial peculiar velocities allows a determination of β (Strauss 1989; Kaiser *et al.* 1991; Hudson 1994). One can make the comparison more sophisticated by smoothing the two velocity fields before comparing them (Davis, Nusser & Willick 1997), or by using the predicted velocity field to minimize the scatter in the distance indicator relation from which the peculiar velocities are measured in the first place (Strauss

1989; Roth 1994; Schlegel 1995; Shaya, Peebles, & Tully 1995; Willick *et al.* 1997b). For *IRAS* galaxies, these velocity comparisons yield values of β_I ranging from 0.49 ± 0.07 (Willick *et al.* 1997b) to 0.86 ± 0.14 (Kaiser *et al.* 1991), depending on the details of the analysis, the smoothing scale, and the data used. Davis *et al.* (1997) have claimed inconsistencies between the peculiar velocity and redshift survey data in the context of GI and linear biasing.

Alternatively, one can use the POTENT method (§ 2) to recover the density fluctuation field from the peculiar velocity data and use equation (3), or its nonlinear extension (see the discussion in § 2.4), to compare to the galaxy density field. Dekel *et al.* (1993, hereafter PI93) carried out such an analysis, using the *IRAS* 1.936 Jy redshift survey (Strauss *et al.* 1992b), and the Mark II compilation of peculiar velocities (Burstein 1989). The advantages of the differential form include the facts that the direct comparison of densities is *local* (whereas the velocity field is sensitive to the mass distribution in a large volume, perhaps even outside the sampled volume), it is independent of reference frame, and it allows direct control over the smoothing of the fields. Monte-Carlo tests showed that the POTENT density field of PI93 was biased in a variety of ways, forcing the use of an elaborate maximum likelihood technique to quantify the consistency between data and model, and to measure β_I. PI93 did find consistency, and concluded that $\beta_I = 1.28^{+0.75}_{-0.59}$ at 95% confidence. A similar comparison of POTENT densities with optically selected galaxies by Hudson *et al.* (1995) yielded an acceptable fit, with $\beta_{opt} = 0.74 \pm 0.13$ (1σ).

The current paper, like PI93, follows the general approach of a *density* comparison, but is a significant step forward due to a number of improvements in the quantity and quality of the data and the methods of analysis. In particular:

- The current analysis is based on the Mark III catalog of peculiar velocities (Willick *et al.* 1995, 1996, 1997a). With ~ 3400 galaxies, it is more than three times bigger than the Mark II data set, and has better space coverage. The data sets composing the Mark III catalog have been treated with more care, especially in self-consistently calibrating the Tully-Fisher (TF) relations, grouping, and correcting for inhomogeneous Malmquist bias (§ 2).

- The POTENT method for deriving the density fluctuation field δ_P from peculiar velocity data has been much improved since PI93 (§ 2; Dekel *et al.* 1997).

- The current analysis uses the *IRAS* 1.2 Jy redshift survey (Fisher *et al.* 1995), containing twice as many galaxies as in the 1.936 Jy survey.

- The method for deriving a uniform galaxy density field δ_I from the redshift data has been improved in several ways since PI93 (§ 3).

- The availability of much more realistic simulations of both the *IRAS* and Mark III datasets (§ 4.1; Kolatt *et al.* 1996) allows much better error analysis in both the peculiar velocity and density fields, and in the comparison.

We use these simulations to assess biases in our determination of β_I. Unlike PI93, these biases turn out to be negligible, allowing us to sidestep the rather elaborate likelihood analysis of that paper. Indeed, we will use the simulations themselves as a guide to whether our data are statistically consistent with the null hypothesis of GI and linear biasing. We also use them to quantify the statistical errors in our final results.

The current analysis compares the density fields smoothed with a Gaussian window of radius $12\,h^{-1}$Mpc, where the fluctuations are of order unity and therefore require a mildly nonlinear treatment. The POTENT analysis computes the density field δ_P, a generalization of $-f(\Omega)^{-1}\boldsymbol{\nabla}\cdot\boldsymbol{v}$ that is a nonlinear function of Ω and the spatial partial derivatives of the observed $\boldsymbol{v}(\boldsymbol{x})$ (equation 15). The *IRAS* reconstruction, in turn, yields a mildly nonlinear galaxy density field, δ_I, that is a weak function of Ω and b_I — only via the corrections from redshift space to real space (equations [16] and [17]). Equation (3) is thus replaced by

$$\delta_I = b_I\,\delta_P \ . \tag{4}$$

The Ω dependence of equation (3) is already included in δ_P. The density fields are recovered from the data for assumed values of Ω and b_I. Then, b_I is determined by equation (4), and β_I is quoted. The analysis is carried out for several initial values of Ω and b_I to confirm the robustness of the estimate of β_I.

The present work is less ambitious than PI93 in one respect. An attempt was made in PI93 to use the nonlinear effects to break the degeneracy between b_I and Ω, but with only limited success. The resulting constraints on each parameter separately were quite weak, indicating that the nonlinear effects associated with these data are not sufficient for this purpose. Even with the new data, the nonlinear effects are comparable to the errors that accompany the reconstructions. Furthermore, if we are to consider nonlinear gravity, we should also allow nonlinear extensions to the biasing relation, equation (2), but in practice, it is not clear how to distinguish nonlinear gravity effects from nonlinear biasing effects. We therefore limit ourselves in this paper to determining the degenerate combination β_I.

The outline of this paper is as follows: In § 2 we discuss the POTENT reconstruction of the mass-density field from peculiar velocities, and in § 3 the reconstruction of the galaxy density field from the *IRAS* redshift survey. In § 4 we use mock catalogs to evaluate the random and systematic errors in the reconstructed density fields, and to minimize them if possible. In § 5 we describe our method of comparison of the two fields, and estimate the systematic and random errors in the measurement of β_I. In § 6 we perform the comparison of the real data, evaluate goodness of fit, and determine the value of β_I. In § 7 we conclude with the implications of our results and compare them with other recent determinations of β_I.

2.   POTENT RECONSTRUCTION FROM PECULIAR VELOCITIES

The POTENT procedure recovers the underlying mass-density fluctuation field from a whole-sky sample of observed radial peculiar velocities. The steps involved are: (a) preparing the data for POTENT analysis, including grouping and correcting for Malmquist bias, (b) smoothing the peculiar velocities into a uniformly-smoothed radial velocity field with minimum bias, (c) applying the ansatz of gravitating potential flow to recover the potential and three-dimensional velocity field, and (d) deriving the underlying density field δ_P by an approximation to GI in the mildly nonlinear regime. The revised POTENT method, which grew out of the original method of Dekel, Bertschinger & Faber (1990, DBF), is described in detail in Dekel *et al.* (1997) and is reviewed in the context of other methods by Dekel (1997a, 1997b). We emphasize below the important improvements since PI93.

2.1.   Mark III Data

We use the most comprehensive catalog of peculiar velocity data available today, the Mark III catalog (Willick *et al.* 1995, 1996, 1997a), which is a careful compilation of several data sets of spiral and elliptical galaxies under the assumption that all galaxies trace the same underlying velocity field. The non-trivial procedure of merging the data sets accounts for differences in the selection criteria, the quantities measured, the method of measurement and the TF calibration techniques. The data per galaxy consist of a redshift z, a "forward" TF (or $D_n - \sigma$) inferred distance, d, and an error. The radial peculiar velocity is then $u = cz - d$.

The Mark II data used in PI93 consisted of about 1000 galaxies (mostly ellipticals by Lynden-Bell *et al.* 1988 and spirals by Aaronson *et al.* 1982). The extended Mark III catalog consists of ~ 3400 galaxies, including the Mark II data (with improvements to the uniformity of Aaronson *et al.* 1982 by Tormen & Burstein 1995) as well as newer spiral data sets by Courteau (1992), Willick (1991), Mathewson, Ford, & Buchhorn (1992), and Han & Mould (1992, and references therein). This sample enables a reasonable recovery of the dynamical fields with $12 \, h^{-1} \mathrm{Mpc}$ smoothing in a sphere of radius $\sim 60 \, h^{-1} \mathrm{Mpc}$ about the Local Group (LG).

2.2.   Correcting Malmquist Bias

Even after the selection bias in the TF calibration is properly corrected, the inferred distance d for a given galaxy, and therefore its peculiar velocity, suffer from a Malmquist bias due to the cross-talk between the random distance errors and the distribution of galaxies about d. We correct this bias in a statistical way before using the velocities as input to POTENT. If the scatter around the TF relation is Gaussian (Willick *et al.* 1997a), i.e., the absolute magnitude M is distributed normally for a given log-linewidth η with standard

deviation σ_m, then the forward inferred distance of a galaxy at a true distance r is distributed log-normally about r, with relative error $\Delta \approx 0.46\sigma_m$. Given d, the expectation value of r is

$$E(r|d) = \frac{\int_0^\infty r P(r|d)dr}{\int_0^\infty P(r|d)dr} = \frac{\int_0^\infty r^3 n(r) \exp\left(-\frac{[\ln(r/d)]^2}{2\Delta^2}\right) dr}{\int_0^\infty r^2 n(r) \exp\left(-\frac{[\ln(r/d)]^2}{2\Delta^2}\right) dr} \ , \tag{5}$$

where $n(r)$ is the number density in the underlying distribution from which the galaxies were selected. The deviation of $E(r|d)$ from d is the bias. The homogeneous part ($n = $ constant), which arises from the three-dimensionality of space, reduces to $E(r|d) = d\,e^{3.5\Delta^2}$, and thus is easy to correct.

Fluctuations in $n(r)$ are responsible for the inhomogeneous Malmquist bias (IM), which, if uncorrected, will systematically enhance the inferred density perturbations δ_P, and thus the deduced value of β. The galaxies are first grouped in z-space (Willick *et al.* 1996), reducing the distance error of each group and thus weakening the bias. The inferred distance of each object is then replaced by $E(r|d)$, with an assumed $n(r)$ properly corrected for grouping and for redshift cutoffs in the data. We use high-resolution density fields of *IRAS* and optical galaxies (Hudson 1993) for the required input $n(r)$ for spirals and ellipticals respectively. Tests with mock data demonstrated that the IM bias can be reduced to the level of a few percent in δ_P (Dekel *et al.* 1997; Eldar, Dekel & Willick 1997).

The resultant correction to δ_P is less than 20% even at the highest peaks, and is much smaller throughout most of space. Thus, the use of *IRAS* data for the IM correction of the Mark III data introduces only a weak coupling to the POTENT output density field. For the purpose of this paper, the POTENT and *IRAS* density fields can thus be regarded as essentially independent data sets. Indeed, a different IM correction procedure in the framework of an "inverse" TF analysis, which does not rely on external input for $n(r)$, yields a density field that is consistent with the result of the forward analysis applied here (Dekel *et al.* 1997; Eldar, Dekel & Willick 1997; following Landy & Szalay 1992).

2.3. Smoothing the Radial Velocities

The goal of the POTENT analysis is to use the set of radial peculiar velocities u_i at positions $\boldsymbol{d}_i$ to determine the underlying continuous fields $\boldsymbol{v}(\boldsymbol{x})$ and $\delta(\boldsymbol{x})$, both smoothed with a Gaussian of radius R_s. We denote hereafter a 3D Gaussian smoothing window of radius R by GR, such that $12\,h^{-1}$Mpc smoothing is indicated by G12, and so on. The first, most difficult step is the smoothing, or interpolation, into a continuous radial velocity field $u(\boldsymbol{x})$. The goal is minimum bias compared to the radial component of the true field, had it been sampled noiselessly, densely and uniformly, and perfectly smoothed in 3D. The radial velocity field at $\boldsymbol{x}_c$ is taken to be the value of an appropriate *local* velocity model $\boldsymbol{v}(\alpha_k, \boldsymbol{x} - \boldsymbol{x}_c)$ at $\boldsymbol{x} = \boldsymbol{x}_c$. The model parameters α_k are obtained by minimizing the weighted

sum of residuals,

$$\sum_i W_i \left[u_i - \hat{\boldsymbol{x}}_i \cdot \boldsymbol{v}(\alpha_k, \boldsymbol{x}_i) \right]^2 , \tag{6}$$

within an appropriate local window $W_i \equiv W(\boldsymbol{x}_i, \boldsymbol{x}_c)$. The window is a Gaussian, modified such that it minimizes the combined effect of the following three types of errors.

Tensor window bias. Because the radial peculiar velocity vectors are not parallel, the u_i's cannot be averaged as scalars, and $u(\boldsymbol{x}_c)$ requires a fit of a local 3D velocity model. The POTENT version of DBF used the simplest local model with three parameters,

$$\boldsymbol{v}(\boldsymbol{x}) = \boldsymbol{B} , \tag{7}$$

for which the solution can be expressed explicitly in terms of a tensor window function. However, a bias occurs because the tensorial correction to the spherical window has conical symmetry, and because variations of $\boldsymbol{v}$ within the window, such as a spherical infall pattern, may be interpreted erroneously as a local bulk velocity. This bias amounts to $\sim 300\,\mathrm{km\,s^{-1}}$ at the Great Attractor (see below). The window bias can be reduced by introducing shear into the model,

$$\boldsymbol{v}(\boldsymbol{x}) = \boldsymbol{B} + \bar{\bar{\boldsymbol{L}}} \cdot (\boldsymbol{x} - \boldsymbol{x}_c) , \tag{8}$$

where $\bar{\bar{\boldsymbol{L}}}$ is a symmetric tensor that automatically ensures local irrotationality. The additional six terms tend to "absorb" most of the bias, leaving $\boldsymbol{v}(\boldsymbol{x}_c) = \boldsymbol{B}$ less biased. Unfortunately, a high-order model tends to pick undesired small-scale noise. The optimal compromise for the Mark III data was found to be equation (8) out to $r = 40\,h^{-1}\mathrm{Mpc}$, smoothly changing to equation (7) beyond $60\,h^{-1}\mathrm{Mpc}$ (Dekel *et al.* 1997).

Sampling-gradient bias. Gradients in the true velocity field within the effective window, coupled with the non-uniform sampling, will introduce a bias if the smoothing is not equal-volume weighted. We thus weight each object in proportion to the local volume that it "occupies", estimated by $V_i \propto R_{4,i}^3$, where R_4 is the distance to the 4-th nearest object. Simulations show that this reduces the bias to negligible levels for $r < 60\,h^{-1}\mathrm{Mpc}$, except for the very sparsely sampled Galactic zone of avoidance. We will use the $R_4(\boldsymbol{x})$ field as a flag for poorly sampled regions, to be excluded from the comparison with *IRAS*.

Random errors. The ideal weighting for reducing the effect of Gaussian noise would be $W_i \propto \sigma_i^{-2}$, where σ_i are the distance errors. Unfortunately, this weighting spoils the volume weighting. As a compromise we weight by both, *i.e.*

$$W(\boldsymbol{x}_i, \boldsymbol{x}_c) \propto V_i\,\sigma_i^{-2}\,\exp[-(\boldsymbol{x}_i - \boldsymbol{x}_c)^2/2R_s^2] . \tag{9}$$

2.4. From Radial Velocity to Density Field

Under GI, the large-scale velocity field is expected to be irrotational, $\boldsymbol{\nabla} \times \boldsymbol{v} = 0$. This remains a good approximation in the mildly nonlinear regime as long as the field is properly

smoothed (Bertschinger & Dekel 1989; DBF). This implies that the velocity field can be derived from a scalar potential, $\boldsymbol{v}(\boldsymbol{x}) = -\boldsymbol{\nabla}\Phi(\boldsymbol{x})$, and thus the potential can be computed by integration along the lines of sight,

$$\Phi(\boldsymbol{x}) = -\int_0^r u(r', \theta, \phi)dr' . \tag{10}$$

The two missing transverse velocity components are then recovered by differentiation.

The subsequent derivation of the underlying mass-density fluctuation field requires a solution to the equations of GI in the mildly nonlinear regime. The linear approximation is limited to the small dynamical range between a few tens of megaparsecs and the $\sim 100\,h^{-1}$Mpc extent of the current samples. Our current peculiar velocity samples enable us to perform reliable dynamical analyses with a smoothing radius as small as $\sim 10\,h^{-1}$Mpc, where $|\boldsymbol{\nabla}\cdot\boldsymbol{v}|$ reaches values larger than unity and therefore nonlinear effects play a role. We appeal to the Zel'dovich (1970) approximation, which is known to be a successful tool in the mildly nonlinear regime. Substituting an Eulerian version of the Zel'dovich approximation in the continuity equation yields (Nusser et $al.$ 1991)

$$\delta_c(\boldsymbol{x}) = \|\mathbf{I} - f^{-1}\partial\boldsymbol{v}/\partial\boldsymbol{x}\| - 1 , \tag{11}$$

where the bars denote the Jacobian determinant and $\mathbf{I}$ is the unit matrix. This was the approximation used in PI93. Tests with N-body simulations show that it does an excellent job for $\delta \geq 0$, but it tends to be an overestimate for $\delta < 0$ (Mancinelli et $al.$ 1994; Ganon et $al.$ 1997).

The Zel'dovich displacement is first order in f^{-1} and $\boldsymbol{v}$, and therefore the determinant in δ_c includes second- and third-order terms as well, involving sums of double and triple products of partial derivatives:

$$\delta_c = -f^{-1}\boldsymbol{\nabla}\cdot\boldsymbol{v} + f^{-2}\Delta_2 + f^{-3}\Delta_3 , \tag{12}$$

where

$$\Delta_2(\boldsymbol{x}) = \sum_{i<j}\left[\left(\frac{\partial v_i}{\partial x_j}\right)^2 - \frac{\partial v_i}{\partial x_i}\frac{\partial v_j}{\partial x_j}\right] \tag{13}$$

and

$$\Delta_3(\boldsymbol{x}) = \sum_{i,j,k}\left[\frac{\partial v_i}{\partial x_i}\frac{\partial v_j}{\partial x_k}\frac{\partial v_k}{\partial x_j} - \frac{\partial v_1}{\partial x_i}\frac{\partial v_2}{\partial x_j}\frac{\partial v_3}{\partial x_k}\right], \tag{14}$$

where the sum is over the three cyclic permutations of $(i, j, k) = (1, 2, 3)$. The approximation can be improved by slight adjustments to the coefficients of the three terms in equation (12),

$$\delta_{c+} = -(1 + \epsilon_1)f^{-1}\boldsymbol{\nabla}\cdot\boldsymbol{v} + (1 + \epsilon_2)f^{-2}\Delta_2 + (1 + \epsilon_3)f^{-3}\Delta_3 . \tag{15}$$

These coefficients were empirically tuned to best fit a family of CDM simulations of $12\,h^{-1}$Mpc smoothing over the whole range of δ values, with $\epsilon_1 = 0.06$, $\epsilon_2 = -0.13$ and $\epsilon_3 = -0.3$. This approximation is found to be robust to unknown quantities such as the value of Ω, the shape of the power spectrum, and the degree of nonlinearity as determined by the fluctuation amplitude and the smoothing scale (Ganon et $al.$ 1997). We adopt $\delta_P = \delta_{c+}$, equation (15), in this paper.

3.   THE *IRAS* RECONSTRUCTION

The redshifts of galaxies in the *IRAS* sample differ from the true distances by the same peculiar velocities that one is attempting to measure in the Mark III dataset,

$$cz = r + \hat{\boldsymbol{x}} \cdot \boldsymbol{v}(\boldsymbol{x}) \; , \tag{16}$$

where cz and $\boldsymbol{v}$ are measured in the same frame of reference. Because of peculiar velocities, the galaxy density field measured in redshift space, $\delta_g(\boldsymbol{s})$, differs systematically from that in real space, $\delta_g(\boldsymbol{x})$ (Kaiser 1987; see reviews in Dekel 1994; Strauss & Willick 1995; Strauss 1997b). Gravitational instability theory enables us to correct for the effects of these velocities. Given the divergence field $\boldsymbol{\nabla} \cdot \boldsymbol{v}$ and appropriate boundary conditions, the velocity is

$$\boldsymbol{v}(\boldsymbol{x}) = -\frac{1}{4\pi} \int \boldsymbol{\nabla} \cdot \boldsymbol{v}(\boldsymbol{x}') \, \frac{\boldsymbol{x}' - \boldsymbol{x}}{|\boldsymbol{x}' - \boldsymbol{x}|^3} \, d^3\boldsymbol{x}' \; . \tag{17}$$

In the linear regime, equation (3) yields

$$\boldsymbol{v}(\boldsymbol{x}) = \frac{\beta}{4\pi} \int \delta_g(\boldsymbol{x}') \, \frac{\boldsymbol{x}' - \boldsymbol{x}}{|\boldsymbol{x}' - \boldsymbol{x}|^3} \, d^3\boldsymbol{x}' \; . \tag{18}$$

Of course, the input density field is defined in real space, which requires that we find a simultaneous solution for the real space density and velocity fields. Yahil *et al.* (1991, hereafter YSDH) and Strauss *et al.* (1992a) describe an iterative technique which implements equations (16) and (18) to calculate the peculiar velocity field and convert the redshifts to distances for an assumed value of β. Our current implementation of the iteration procedure has three new features: the treatment of triple-valued zones, the filtering of the density field, and the nonlinear corrections.

3.1.   Modeling Triple-Valued Zones

Given a model for the velocity field from equations (3) and (17), one can compute the redshift-distance relation along a line of sight to a given galaxy from equation (16). This can be compared with the observed redshift of the galaxy to solve for the distance r. However, in the vicinity of prominent overdensities, the redshift as a function of distance can become non-monotonic, such that there are three different distances corresponding to a given redshift. YSDH describe a distance-averaging procedure that recognizes the existence of triple-valued zones. The more general approach applied here is inspired by the VELMOD maximum-likelihood analysis of Willick *et al.* (1997b). Along a given line of sight, we ask for the joint probability distribution of observing a galaxy with redshift cz, flux density $\mathcal{F}$ and (unknown) distance r:

$$P(cz, \mathcal{F}, r) = P(cz|r) \times P(\mathcal{F}|r) \times P(r) \quad . \tag{19}$$

The first term is provided by our velocity field model. We assume that the contributions to the velocity field of shot noise, the nonlinearities, and the gravitational influence of material within one smoothing length are incoherent, and thus we model the scatter around the assumed redshift-distance relation as a Gaussian, with dispersion σ_v:

$$P(cz|r) = \frac{1}{\sqrt{2\,\pi}\sigma_v} \exp\left(-\frac{[cz - \hat{\boldsymbol{x}} \cdot \boldsymbol{v}(\boldsymbol{x})]^2}{2\,\sigma_v^2}\right) . \tag{20}$$

For what follows, we set $\sigma_v = 150\,\mathrm{km\,s^{-1}}$, independent of position (cf., Willick *et al.* 1997b). Like YSDH, we collapse clusters of galaxies to a common redshift.

The second term of equation (19) is given by the luminosity function of galaxies, $\Phi(L)$,

$$P(\mathcal{F}|r) = \Phi(L{=}4\,\pi r^2 \mathcal{F})dL/d\mathcal{F} \propto r^2 \Phi(L) , \tag{21}$$

where the derivative is needed because the probability density is defined in terms of $\mathcal{F}$, not L.

The third term in equation (19) is given by the galaxy density distribution along the line of sight:

$$P(r) \propto n(\boldsymbol{x})r^2 \propto [1 + \delta_g(\boldsymbol{x})]\,r^2 . \tag{22}$$

Finally, given the joint probability distribution, equation (19), we estimate the distance of a given galaxy to be the expectation value,

$$\langle r \rangle = \frac{\int r P(cz, \mathcal{F}, r)\,dr}{\int P(cz, \mathcal{F}, r)\,dr} . \tag{23}$$

3.2. Power-Preserving Filtering

In a flux-limited sample, the mean number density of objects is a monotonically decreasing function of distance, and therefore the shot noise in the density field and predicted peculiar velocity field both increase with distance. Several approaches to this problem have been taken in the literature, but they all involve variable smoothing, which is not desirable for the purpose of comparing with the uniformly-smoothed POTENT output. YSDH set a top-hat smoothing length equal to the mean inter-particle spacing, which of course increases with distance from the origin. Nusser & Davis (1995) expand the density field in spherical harmonics and radial Bessel functions. Given a finite number of angular modes (independent of distance) the effective smoothing length again increases with scale. Fisher *et al.* (1995) perform a similar expansion of the density field, but in addition, they apply a Wiener filter to the expansion coefficients.

The Wiener filter (cf., Press *et al.* 1992) provides the *minimum variance* reconstruction of the smoothed density field, given *a priori* knowledge of the underlying power spectrum

and the noise (cf., Lahav *et al.* 1994; Zaroubi *et al.* 1995). Let the observed smoothed density field be $\delta_O(\boldsymbol{x})$, and the true underlying density field $\delta_T(\boldsymbol{x})$, such that $\delta_O(\boldsymbol{x}) = \delta_T(\boldsymbol{x}) + \epsilon(\boldsymbol{x})$, where ϵ is the local contribution from shot noise. We wish to minimize the squared error in the density field by filtering the observed density field. Filtering is best done in Fourier space, $\tilde{\delta}_F(\boldsymbol{k}) = F(\boldsymbol{k})\tilde{\delta}_O$, where the tilde indicates Fourier Transform. Parseval's Theorem allows us to write the squared error of the filtered density field as

$$\langle [\delta_T(\boldsymbol{x}) - \delta_F(\boldsymbol{x})]^2 \rangle = \langle [\tilde{\delta}_T(\boldsymbol{k}) - F(\boldsymbol{k})\tilde{\delta}_O(\boldsymbol{k})]^2 \rangle \; . \tag{24}$$

Minimizing this expression with respect to the filter F gives the *Wiener* filter,

$$F_W = \frac{\langle \tilde{\delta}_T^2 \rangle}{\langle \tilde{\delta}_T^2 \rangle + \langle \tilde{\epsilon}^2 \rangle} \; . \tag{25}$$

If the underlying density field and the noise are both Gaussian-distributed, the Wiener filter provides the most probable solution for the mean underlying field.

However, the Wiener filter has a serious drawback for the present application. The expectation value of the square of the filtered field is:

$$\langle F_W^2 \, \tilde{\delta}_O^2 \rangle = \left(\frac{\langle \tilde{\delta}_T^2 \rangle}{\langle \tilde{\delta}_T^2 \rangle + \langle \tilde{\epsilon}^2 \rangle} \right)^2 \left(\langle \tilde{\delta}_T^2 \rangle + \langle \tilde{\epsilon}^2 \rangle \right) = F_W \langle \tilde{\delta}_T^2 \rangle \; . \tag{26}$$

Since $F_W < 1$, the Wiener-filtered field has a variance that is always *smaller* than that of the true underlying field (contrast this with the case of the observed field, whose variance is systematically *greater* than that of the true field, by an amount $\langle \tilde{\epsilon}^2 \rangle$). Even worse, because ϵ is in general an increasing function of distance in a flux-limited sample, the variance in the Wiener-filtered field is a decreasing function of distance. In fact, when the noise dominates over the signal, the most likely value of δ tends to the mean value of the prior model, *i.e.*, $\delta \to 0$. This is unacceptable for the current application, because any such systematic effect in δ_g will translate directly into a bias in β_I (cf., equation [3]).

With this in mind, Yahil (1994) developed a *power preserving* variant of the Wiener filter, defined by

$$F_Y = F_W^{1/2} \; . \tag{27}$$

A calculation analogous to that of equation (26) shows that $\langle F_Y^2 \, \tilde{\delta}_O^2 \rangle = \langle \tilde{\delta}_T^2 \rangle$; that is, the filter preserves the power in the underlying true density field. It is not optimal in the sense of minimum variance, but it is straightforward to show that the ratio of the mean square difference between true and filtered fields for the Yahil and Wiener filters can be expressed in terms of the signal-to-noise ratio as

$$\frac{\langle (\tilde{\delta}_T - F_Y \, \tilde{\delta})^2 \rangle}{\langle (\tilde{\delta}_T - F_W \, \tilde{\delta})^2 \rangle} = 1 + \left[(1+y)^{1/2} - y^{1/2} \right]^2 \; , \quad y \equiv \frac{\langle \delta_T^2 \rangle}{\langle \epsilon^2 \rangle} \; . \tag{28}$$

The square root of this ratio approaches unity when the signal dominates, $y \gg 1$. It rises to only 1.08 when $y = 1$, and is still only 1.41 even when the noise dominates, $y \ll 1$. That is,

we pay only a mild price in density field errors when replacing the Wiener filter by the Yahil filter, while gaining an unbiased estimate of the variance of the density field everywhere.

Applying any of these filters requires that we first define a continuous density field. We assign the galaxies in the sample (weighted by the inverse of the selection function) via cloud-in-cell to a 128^3 Cartesian grid of spacing $2\,h^{-1}$Mpc, and smooth the density field further with a small Gaussian window of G5 (or G4, see § 4.3). We then apply the Yahil filter to the Fourier Transform of the density field, where $\langle\delta_T^2\rangle$ and $\langle\epsilon^2\rangle$ are calculated as appropriate for the smoothing window, using as a prior the power spectrum of $IRAS$ galaxies as derived by Fisher $et\ al.$ (1993).

Of course, the appropriate Yahil filter to apply at each point in real space is a function of the distance from the origin, as the noise term $\langle\epsilon^2\rangle$ is an increasing function of distance. We thus calculate a series of Yahil-filtered density fields, each assuming the shot noise appropriate for one of 13 different distances from the origin: 5, and 20 to $240\,h^{-1}$Mpc in steps of $20\,h^{-1}$Mpc. The value of the Yahil-filtered density field at any other point in space is then spline-interpolated from those 13 fields.

3.3. Nonlinear Correction

Because we are calculating a continuous density field, we can take advantage of nonlinear extensions to equation (3), before plugging $\boldsymbol{\nabla}\!\cdot\!\boldsymbol{v}$ into equation (17) and solving it in Fourier space. Nusser $et\ al.$ (1991) provided a simple functional fit to the inverse of equation (11), which does a good job of fitting nonlinear effects for smoothing scales of G5 and larger (see also Mancinelli & Yahil 1995). The Nusser $et\ al.$ expression has a mean differing from zero, which gives rise to an erroneous monopole term in the derived velocity field. Therefore, we use a generalized expression whose mean vanishes to second order:

$$\boldsymbol{\nabla}\!\cdot\!\boldsymbol{v} = -\beta\,\frac{(1+\alpha^2\sigma_g^2)\,\delta_g + \alpha b\sigma_g^2}{1+\alpha b^{-1}\delta_g}\ ,\tag{29}$$

where $\sigma_g^2 \equiv \langle\delta_g^2\rangle$, and α is a constant to be determined empirically. The best fit found in CDM N-body simulations is $\alpha = 0.28$ (Ganon $et\ al.$ 1997).

The iteration technique used to find self-consistent density and velocity fields for the $IRAS$ sample is similar to that described in Strauss $et\ al.$ (1992a), with the additional application of the Yahil filter, the nonlinear correction of equation (29), and the refined method of § 3.1 for determining the distance of each galaxy. We do all the calculations in the Local Group frame, on a 128^3 grid centered on the observer and oriented along Supergalactic coordinates. The density field is calculated within a sphere of radius $128\,h^{-1}$Mpc, and is assumed to be uniform outside this sphere and inside the bounding cube. Periodic boundary conditions are assumed. We term the resulting density field of $IRAS$ galaxies in real space δ_I, for specific assumed values of Ω and b_I.

4. EVALUATING THE RECONSTRUCTION METHODS

The POTENT and *IRAS* reconstruction methods laid out in the previous sections are evaluated using mock catalogs based on *N*-body simulations. The systematic errors are investigated and corrected as much as possible, and the random statistical errors are quantified. We first describe the mock catalogs (§ 4.1), then summarize the testing of the POTENT and the *IRAS* methods themselves (§ 4.2, § 4.3), and eventually describe the calibration of the comparison (§ 5.2).

4.1. The Mock Catalogs

The mock catalogs and the underlying *N*-body simulations are described in detail in Kolatt *et al.* (1996); we present only a brief outline here.

A special effort was made to generate simulations that mimic the actual large-scale structure in the real universe, in order to take into account any possible dependence of the errors on the signal. The present-day density field, G5 smoothed, is taken to be that of *IRAS* galaxies as reconstructed by the method described in § 3, with $b_I = \Omega = 1$. The field is traced back in time to remove nonlinear effects by integrating the Zel'dovich-Bernoulli equation (Nusser & Dekel 1992). Non-Gaussian features are removed, and structure on scales smaller than the smoothing length is added using the method of constrained realizations (Hoffman & Ribak 1991), with the power spectrum of the *IRAS* 1.2 Jy survey (Fisher *et al.* 1993) as a prior. The resulting density field is fed as initial conditions to a PM *N*-body code (Bertschinger & Gelb 1991) which then follows the nonlinear evolution under gravity, with $\Omega = 1$. The present epoch is defined by an *rms* density fluctuation of $\sigma_8 = 0.7$ at a top-hat smoothing of radius $8\,h^{-1}\mathrm{Mpc}$, as is observed for *IRAS* galaxies (Fisher *et al.* 1994a). The periodic box of side $256\,h^{-1}\mathrm{Mpc}$ is simulated with a 128^3 grid and 128^3 particles.

Next, "galaxies" are identified in the simulation and assigned the relevant physical properties. They are then "observed" to make mock catalogs that include all the relevant errors and selection effects in both the peculiar velocity and redshift survey data. For the mock Mark III catalogs, each of the *N*-body particles is considered a galaxy candidate, and is identified as elliptical or spiral, depending on the local neighborhood density of particles (following Dressler 1980). Rich clusters are identified, mimicking the cluster samples in the real data, and the remaining particles are left as candidates for field galaxies. The galaxies are assigned log-linewidths η drawn at random from the observed η distribution function (corresponding to the observed galaxy luminosity function). A Tully-Fisher (or $D_n - \sigma$) relation is assumed, and absolute magnitudes M are randomly scattered about the TF value, $M_{TF}(\eta)$, following a Gaussian distribution of width appropriate to the corresponding sub-sample of the Mark III catalog. Field galaxies are selected in the angular regions corresponding to each of the sub-samples, with the appropriate magnitude limits and redshift cutoffs. The only feature of the observational procedure that is not simulated is the calibration of the TF relations for each sample in the Mark III catalog, and the matching of them in their overlap;

the TF relations are assumed to be known perfectly *a priori.*

These data are used to infer TF distances to all the galaxies in the mock catalog. The "observed" redshifts are taken to be the true velocities of the particles in the simulation. Finally, the galaxies selected are grouped using the same code used for the real data, and are corrected for Malmquist bias as in § 2.2, using the galaxy number density profile $n(r)$ as derived from a randomly selected mock *IRAS* catalog (see § 4.3).

4.2.   Errors in the POTENT Reconstruction

The left panel of Figure 1 demonstrates how well POTENT can do with ideal data of dense and uniform sampling and no distance errors. The reconstructed density field, from input that consisted of the exact, G12-smoothed radial velocities, is compared with the true G12 density field of the simulation. The comparison is done at grid points of spacing $5\,h^{-1}$Mpc inside our "standard" comparison volume of effective radius $40\,h^{-1}$Mpc (see below). We see that no bias is introduced by the POTENT procedure itself. The small *rms* scatter of 2.5% reflects the cumulative effects of small deviations from potential flow, scatter in the nonlinear approximation (equation [15]), and numerical errors (compare to the discussion of the nonlinear effects on this smoothing scale in PI93).

Fig. 1.— Systematic errors in the POTENT analysis. The density field recovered by POTENT from the mock data is compared with the "true" G12 density. The comparison is at uniform grid points within a volume of effective radius $40\,h^{-1}$Mpc. Left: The input to POTENT is the true, G12-smoothed radial velocity. Right: The input is noisy and sparsely-sampled mock data. Shown is $\langle\delta_P\rangle$, the POTENT field averaged over 10 random realizations. The sizes of the symbols are inversely proportional to the estimated error σ_P at that grid point. The solid line is the average over the regression lines in the ten realizations, of slope 0.99 ± 0.06. The error is the standard deviation over the realizations. Only half of the points (randomly selected) within the comparison volume are actually plotted.

We execute the POTENT algorithm on each of ten noisy mock realizations of the Mark III catalog, recovering 10 corresponding density fields. The error in the POTENT density field at each point in space, σ_P, is taken to be the *rms* difference over the realizations, between δ_P and δ_T, the true G12-smoothed density field of the mass in the simulation. This scatter includes both systematic and random errors. We evaluate the density field and the errors on a Cartesian grid with $5\,h^{-1}$Mpc spacing. In the well-sampled regions, which extend in Mark III out to $40-60\,h^{-1}$Mpc, the errors are $\sigma_P \approx 0.1-0.3$, but they are much larger in certain regions at large distances. We will use σ_P below to exclude noisy regions in the POTENT-*IRAS* comparison (§ 5.1). In particular, our standard comparison volume is defined by $\sigma_P < 0.3$ and $R_4 < 9.2\,h^{-1}$Mpc. The resulting volume V has an effective radius $R_e = 40\,h^{-1}$Mpc, defined by $V = (4\pi/3)R_e^3$. The *rms* of σ_P in the standard volume is 0.19.

Some part of these errors is systematic. We can quantify this by comparing δ_T to the recovered mean density field, averaged over the mock catalogs at each point. The right panel of Figure 1 makes this comparison at the points of a uniform grid inside the standard volume. The residuals in this scatter plot ($\langle \delta_P \rangle$ *vs.* δ_T) are the local systematic errors. Their *rms* value over the standard volume (and over the realizations) is 0.06. The corresponding *rms* of the random errors (δ_P *vs.* $\langle \delta_P \rangle$) is 0.18. The systematic and random errors add in quadrature to the total error (δ_P *vs.* δ_T), whose *rms* over the realizations at each point, σ_P, is used in the analysis below.

Systematic errors may also be correlated. We quantify this by performing a regression of δ_P on δ_T: $\delta_P = m\,\delta_T + n$, for each realization, and then measuring the scatter around this best-fit line:

$$S' \equiv \frac{1}{N_{\mathrm{grid}}} \sum_i^{N_{\mathrm{grid}}} \frac{[\delta_{Pi} - (m\,\delta_{Ti} + n)]^2}{\sigma_{Pi}^2} \ , \tag{30}$$

which we compute for each realization within the standard volume. This statistic is not necessarily distributed like χ^2 because the local systematic errors may be correlated, but it should still average to $\langle S' \rangle = 1$ as long as the errors of δ_P about δ_T average to zero. The average of S' over the ten realizations is $\langle S' \rangle = 0.94$, with a standard deviation of 0.14 and therefore an error of 0.05 in the mean. The proximity of this average to unity indicates that, indeed, the local errors roughly average to zero.

The main point of Figure 1 is that the global systematic errors in δ_P as a function of δ_T are small. The figure shows no systematic deviations from the $y = x$ line. The average of the slopes m of the regression lines over the ten realizations is 0.99 (compared to the correct answer of 1.0) with a standard deviation of 0.06. This means that the local systematic errors show no significant correlation with the signal, contributing no bias to the comparison with *IRAS*.
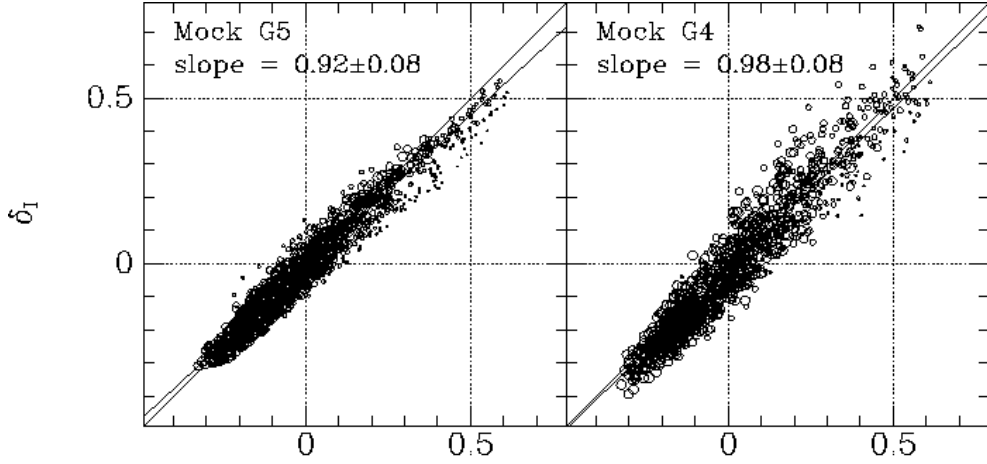
Fig. 2.— Systematic errors in the *IRAS* analysis. The density field reconstructed from the mock data, averaged over 9 realizations, is compared with the "true" G12 density. The comparison is at uniform grid points within the standard volume of effective radius $40\,h^{-1}$Mpc. The line is the average of the regression lines over the realizations. The quoted slope and error are the average and standard deviation over the realizations. Left: with G5 smoothing in the iterative procedure, followed by G10.9 smoothing, showing a 8% (3σ) bias in the slope due to over-smoothing. Right: with G4 smoothing instead of G5 smoothing. Only half of the points (randomly selected) within the comparison volume are actually plotted.

4.3.   Errors in the *IRAS* Reconstruction

Mock *IRAS* redshift surveys are drawn from the simulation following the observed luminosity function and redshift distribution in the *IRAS* 1.2 Jy catalog. We allow each particle to be chosen as a galaxy with equal probability, so $b_I = 1$ for these samples. Those regions of the sky not surveyed in the *IRAS* sample (Strauss *et al.* 1990) are excluded.

Each of nine mock *IRAS* redshift catalogs is now put through the reconstruction code described in § 3, assuming $\Omega = b_I = 1$ and G5 smoothing (but see below). The resulting density fields are smoothed further with a Gaussian of $10.9\,h^{-1}$Mpc to yield a nominal total effective smoothing of G12, in order to match the smoothing used in the POTENT reconstruction. Finally, the density field recovered from each of the nine realizations is compared with the true G12 density field of the simulation.

The average recovered density field is compared with the true field in Figure 2 (left panel). The comparison is done within the standard comparison volume (defined by PO-TENT errors as before). The average slope of the regression lines, shown in this figure, is 0.92 ± 0.08, where the error quoted is the standard deviation over the nine realizations, *i.e.*, the mean slope is about 3σ away from unity. This bias reflects an effective over-smoothing of the *IRAS* density field, which is partly due to the cloud-in-cell algorithm used to place galaxies on the grid. We find that this bias is significantly reduced by artificially replacing

the desired G5 smoothing with a G4 smoothing. After applying the same additional G10.9 smoothing, the average slope of the regression becomes 0.98 ± 0.08, which we regard as satisfactorily unbiased for the purpose of comparison with POTENT. The average density, with G4+G10.9 smoothing, is compared with the true G12 density in the right panel of Figure 2.

The total error in the *IRAS* G12 density field at each grid point, σ_I, is taken to be the *rms* value of $\delta_I - \delta_T$ over the realizations. Its *rms* within the standard volume is $\sigma_I = 0.09$. The random and systematic contributions are estimated to be 0.08 and 0.04 respectively.

As we did with POTENT, we do regressions of each realization of δ_I on δ_T, and quantify the scatter S', following equation (30). We find $\langle S' \rangle = 0.60$ with a standard deviation of ± 0.18. This is significantly smaller than unity, indicating that there are systematic errors in the determination of δ_I that do not average to zero. This is not of major concern; when we later evaluate goodness of fit of the real data using an S'-like statistic, we do not assume that it is distributed like χ^2, but rather determine its distribution directly from the mock realizations.

5. METHOD OF COMPARISON

5.1. Measuring β_I

Since the mock *IRAS* and mock Mark III catalogs are both drawn from the same underlying density field of an $\Omega = 1$ simulation with $b = 1$, a perfect method of comparison should yield $\beta = 1$.

As mentioned in § 4.2, we limit the comparison to the regions where the recovery is reliable in both data sets. In fact, the POTENT errors are typically twice as large as the *IRAS* errors, so we simply restrict ourselves to regions with small σ_P. In addition, in order to minimize the effects of sampling gradient bias, we also avoid regions of sparse sampling as indicated by a large R_4.

We carry out the comparison using three alternative sets of cuts on σ_P and R_4, defining different volumes of space about the Local Group, in order to test our sensitivity to these limits and the volume surveyed. These three cuts are given in the left section of Table 1, which lists the R_4 and σ_P limits, the number of grid points (of spacing $5\,h^{-1}\mathrm{Mpc}$) that satisfy this cut, the effective radius R_e, and the corresponding effective number of independent smoothing volumes N_{eff}, which we estimate by $N_{\mathrm{eff}}^{-1} = N_{\mathrm{grid}}^{-2} \sum_{j=1}^{N_{\mathrm{grid}}} \sum_{i=1}^{N_{\mathrm{grid}}} \exp(-r_{ij}^2/2R_s^2)$, where $R_s = 12\,h^{-1}\mathrm{Mpc}$ is the smoothing radius[7]. This range of comparison volumes repre-

---

[7] This N_{eff} is not simply the ratio of the comparison volume to the effective volume of the smoothing window; it also takes into account the shape of the comparison volume, in the sense that N_{eff} becomes appropriately larger as the shape of the comparison volume deviates from a sphere. Unlike in Hudson *et al.* (1995), N_{eff} is not used in our calculations of errors, and is provided only as a reference.

sents a compromise between our need to avoid noisy and sparsely sampled regions, and our wish to include as large a volume as possible in order to reduce cosmic scatter and come closer to a fair sample. We have already introduced the volume of $R_{\rm e} = 40\,h^{-1}{\rm Mpc}$ as our standard comparison volume.

Our model is that the variables δ_I and δ_P are linear functions of the (unknown) true density field δ_T:

$$\delta_I = b_I \delta_T + \epsilon_I \ , \tag{31}$$

and

$$\delta_P = \delta_T + c + \epsilon_P \ , \tag{32}$$

where ϵ_I and ϵ_P are independent random variables with dispersions σ_I and σ_P, respectively. In this case, we can estimate b_I and c by minimizing the χ^2-like quantity (Lawley & Maxwell 1971):

$$\chi^2 = \sum_i^{N_{\rm grid}} \frac{(\delta_{Pi} - b_I^{-1}\delta_{Ii} - c)^2}{\sigma_{Pi}^2 + b_I^{-2}\sigma_{Ii}^2} \ , \tag{33}$$

with respect to b_I and c, taking into account the errors in both fields. The allowed offset c reflects the zero-point freedom in the TF data and the uncertainty in the mean density in the *IRAS* field. Remember that δ_P already includes the factor $f^{-1}(\Omega)$ (equation 15), which is why β_I does not appear in this equation.

We sample the density fields every $5\,h^{-1}{\rm Mpc}$ while our smoothing length is $12\,h^{-1}{\rm Mpc}$; thus the fields are greatly oversampled, $N_{\rm grid} \gg N_{\rm eff}$. It is meaningful to define a statistic,

$$S = \chi^2/N_{\rm grid} \tag{34}$$

for estimating goodness of fit, as long as we remember that its probability distribution function is not that of a reduced χ^2; the expectation value may be $\langle S \rangle = 1$, but the dispersion reflects $N_{\rm eff}$ and not $N_{\rm grid}$. We thus do not use it to determine the statistical error in b_I^{-1} (as was done for example in Hudson *et al.* 1995). We instead determine the distribution of this statistic from the mock catalogs that satisfy our null hypotheses of GI and linear biasing (§ 5.2), and then compare the value of S obtained from the real data with this distribution (§ 6.2).

5.2. Testing the Comparison with Mock Catalogs

We carry out the comparison using 90 pairs of mock catalogs, pairing the 10 mock POTENT with the 9 mock *IRAS* realizations. For each of the three comparison volumes, we calculate the mean and standard deviation over the pairs of the quantities S and b_I^{-1}, and list them in the middle section of Table 1.

Table 1:

Comparison Volume					Mock Data				Real data	
$\sigma_P <$	$R_4{}^1 <$	N_{grid}	N_{eff}	$R_{\mathrm{e}}{}^1$	$\langle S \rangle$	σ_S	$\langle b^{-1} \rangle$	$\sigma_{b^{-1}}$	S	b_I^{-1}
0.20	8.0	1005	10	31	0.80	0.14	1.04	0.13	1.03	0.83
0.30	9.2	2081	18	40	0.89	0.16	1.03	0.12	1.06	0.89
0.40	10.0	3342	26	46	0.93	0.15	1.02	0.11	1.16	0.93

$^1\, h^{-1}\mathrm{Mpc}$

The most important conclusion from the comparison of the mock data is that our method provides an almost bias-free estimate of b_I^{-1}. This is consistent with our previous tests showing that the POTENT and *IRAS* reconstructions themselves are both hardly biased (§ 4.2, 4.3). When we combine the two, and fit for b^{-1} using equation (33), we find that the values of $\langle b^{-1} \rangle$ deviate by only $2-4\%$ from the correct value of unity.

How significant is this small bias? The 90 pairs of realizations are not all independent because the *IRAS* errors are smaller than the POTENT errors. We crudely estimate the effective number of independent pairs to be ~ 20, so the error in $\langle b^{-1} \rangle$ is roughly $\sigma_{b^{-1}}/\sqrt{20} \sim 0.03$, and thus the apparent 3% bias is not significant.

Another important result is that b^{-1} is fairly robust to the comparison volume in the simulations. A factor of 3 increase in the comparison volume leads to only a 2% change in the value of $\langle b^{-1} \rangle$. One conclusion is that cosmic scatter in the simulations does not play a very important role in the determination of b^{-1} by this method. Our formal error in b^{-1}, the derived scatter in b^{-1} between realizations, does not include the cosmic scatter, and thus underestimates the true error, but not by a large amount. This cosmic scatter will be estimated again similarly in the real universe in § 6 by comparing results from different volumes.

The linear fit of equation (33) is affected by the systematic errors between δ_I and δ_T, and between δ_P and δ_T, that we estimated using equation (30). Given that $\langle S' \rangle$ was substantially less than unity for the *IRAS* reconstruction, we expect σ_I to be somewhat of an overestimate, and therefore, the S of the δ_P-δ_I comparison may be somewhat less than unity. As Table 1 shows, this is indeed what we find for the mock catalogs.

As mentioned above, we will not use S and χ^2 statistics to determine the error in b_I^{-1}. Instead, we use the scatter of b^{-1} over the pairs of mock catalogs as our estimate for the error in the value of b_I^{-1} as derived from the real data. We find that this error is $\pm 0.11 - 0.13$ (middle section of Table 1).

For the offset density we find from the mock catalogs small values of $\langle c \rangle = 0.04$, 0.03 and 0.02 for the three comparison volumes from small to large, with a scatter $\sigma_c = 0.07$, 0.06 and 0.07 respectively. The displacement is thus consistent with zero (although a small displacement is expected, especially at small volumes, due to the fact that the mean density was determined separately in the reconstruction from each *IRAS* mock catalog).
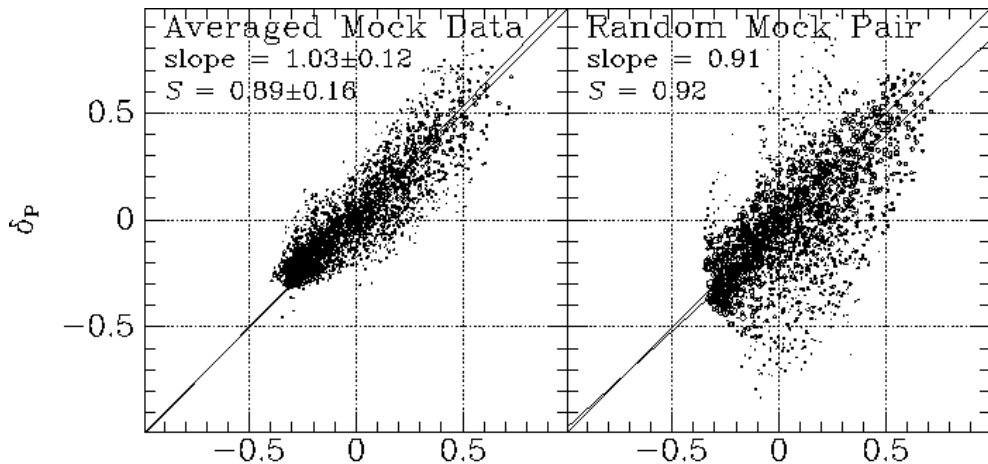
Fig. 3.— POTENT versus IRAS density fields as reconstructed from the mock catalogs within the standard volume. The symbol area is inversely proportional to the error σ_e at each grid point. Left: The average mock POTENT field versus the average mock *IRAS* field. The line slope and the statistic S quoted are the averages over 90 pairs. Right: The density fields in one random pair of realizations.

The left panel of Figure 3 shows a scatter plot of the average of the 10 POTENT realizations versus the average of the 9 *IRAS* realizations, for the standard volume. The line of slope 1.03, which is the average of the slopes over the individual pairs of realizations, is shown for reference (it is not necessarily identical to the best-fit line of $\langle \delta_P \rangle$ versus $\langle \delta_I \rangle$). This figure provides a visual impression of the correlation, in the spirit of Figure 1 and Figure 2. The fact that the slope is close to unity reflects the lack of global bias, and the scatter is the local bias at the grid points.

For a visualization of the actual scatter that one may expect to see in the real universe, we show in the right panel of Figure 3 the comparison of one arbitrary pair of realizations. The scatter is naturally larger than in the left panel, but the correlation is still very strong.

Figure 4 displays contour maps from the mock catalogs; the contouring is described in the figure caption. The upper and middle rows show G12 density fluctuation fields in the Supergalactic plane. The upper row is the average over the realizations of POTENT and *IRAS* in comparison with the true G12 field, and the middle row shows one random realization of each and the difference between them, in units of σ_e, defined in equation (35) below. It is interesting to compare these maps to Figure 6 below, which shows the reconstructions of the real fields. Note that the *IRAS* errors are smaller than the POTENT errors, so the mock *IRAS* reconstruction resembles the true one more closely than the POTENT reconstruction. The difference between the *IRAS* and POTENT fields can be very large, especially at large distances, where we know the POTENT reconstruction is less reliable. For most of the volume, the difference is of low statistical significance.

The bottom row of Figure 4 shows the boundaries of the small ($R_e = 31 \, h^{-1} \mathrm{Mpc}$) and
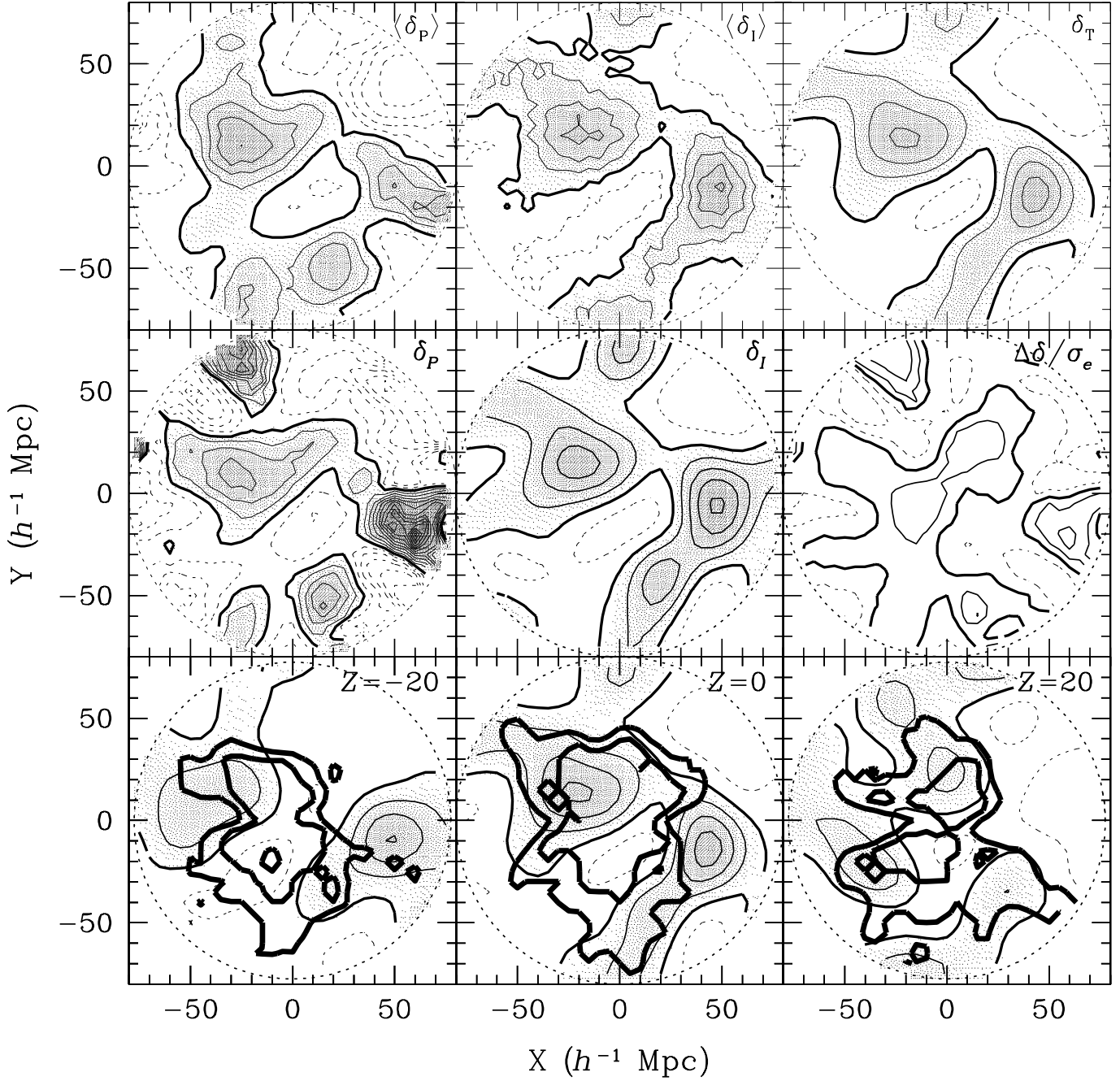
Fig. 4.— Density fluctuation fields of POTENT mass versus *IRAS* galaxies from the mock catalogs. The smoothing is G12. Contour spacing is 0.2 in δ, the heavier contour is $\delta = 0$, solid contours refer to $\delta > 0$ and dashed contours to $\delta < 0$. The density is also indicated by shading. Top row: maps in the Supergalactic plane of the average fields over the mock realizations, in comparison with the true G12 density field of the simulation. Middle row: the density fields reconstructed from one pair of realizations, and the difference field $\Delta\delta$ in units of the error σ_e, with contour spacing of unity. Bottom row: the boundaries of the $R_e = 31$ and $46\,h^{-1}$Mpc comparison volumes in the Supergalactic plane ($Z = 0$) and in two parallel planes ($Z = \pm 20\,h^{-1}$Mpc).

large ($R_{\mathrm{e}} = 46\,h^{-1}$Mpc) comparison volumes in the Supergalactic plane ($Z = 0$) and two parallel planes ($Z = \pm 20\,h^{-1}$Mpc), superposed on the true field. The narrowing of the comparison volume in the Zone of Avoidance (roughly corresponding to $Y = 0$), due to the large R_4, is clearly visible.

6. RESULTS

With the method calibrated using the mock catalogs, we are in a position to compare the POTENT and *IRAS* reconstructions of the real data. We carry out the comparison in the three comparison volumes of Table 1, using the errors σ_P and σ_I as estimated from the mock catalogs in the χ^2 expression of equation (33). We start with a visual comparison using maps (§ 6.1), proceed to a more quantitative analysis of the residuals in order to evaluate the goodness of fit (§ 6.2), and conclude with a measurement of β_I (§ 6.3).

6.1. Visual Comparison

Fig. 5.— POTENT mass ($\Omega = 1$) versus *IRAS* galaxy density fields in the Supergalactic plane, both smoothed G12. The height of the surface plot is proportional to δ. The LG is at the center, GA on the left, PP on the right, and the Sculptor void in between. The region shown is $160\,h^{-1}$Mpc on a side.

Figure 5 shows the G12 density fields δ_P ($\Omega = 1$) and δ_I in the Supergalactic plane. The general correlation is evident — the Great Attractor (GA), Perseus-Pisces (PP), and the void in between, all show up both as dynamical entities and as structures in the galaxy distribution.

More quantitative visual comparisons of the two density fields are provided in Figure 6, which shows in three rows three different planes in Supergalactic coordinates: the Supergalactic plane $Z = 0$, and the planes $Z = -20\,h^{-1}$Mpc and $Z = +20\,h^{-1}$Mpc. The first and second columns show the G12 density fields of POTENT ($\Omega = 1$) and *IRAS*, and the third
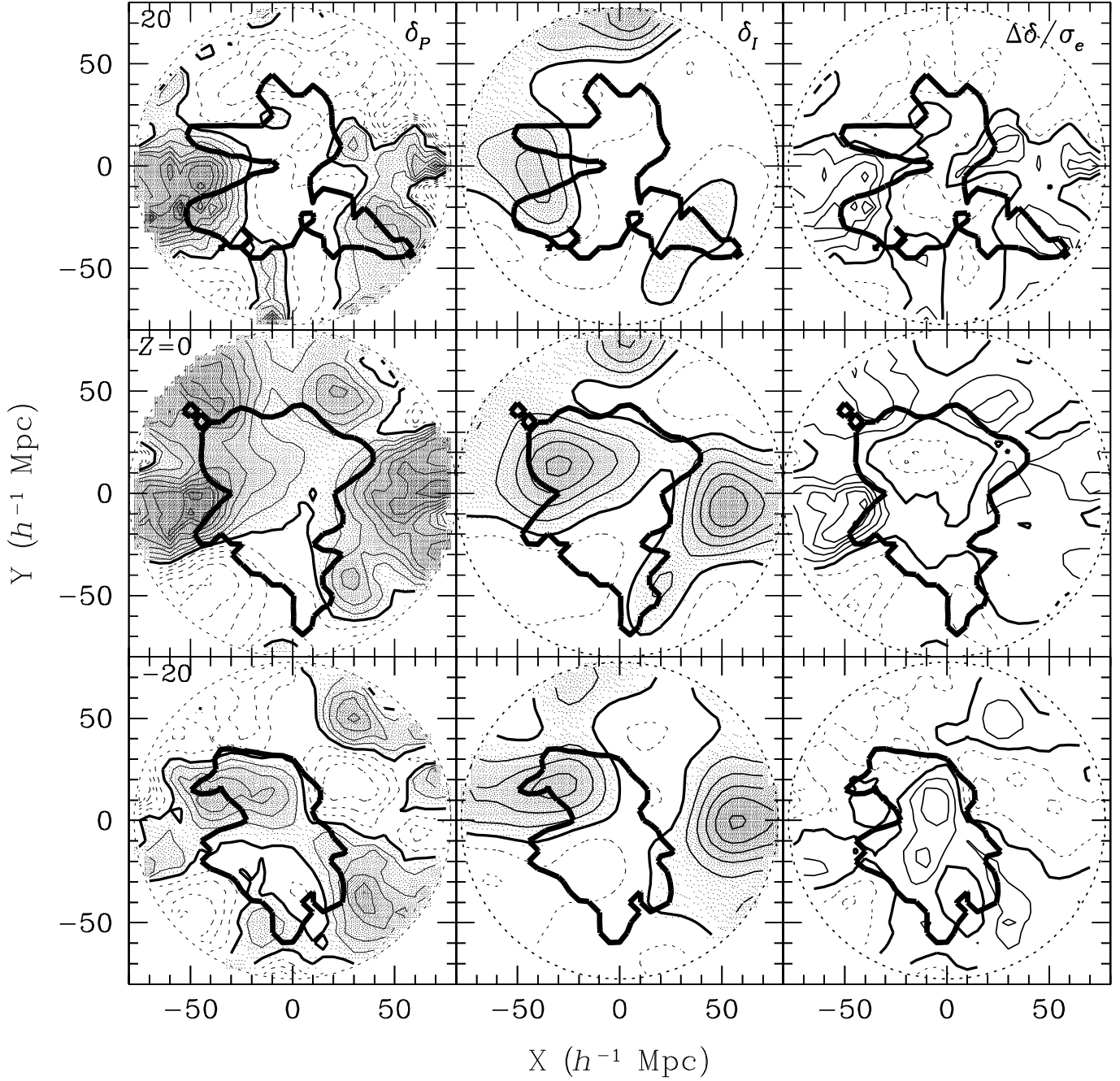
Fig. 6.— Density fluctuation fields of POTENT mass (left-hand column, $\Omega = 1$) versus *IRAS* galaxies (middle column), both smoothed G12. Each row refers to a different slice: $Z = +20 \, h^{-1}$Mpc (top row), $Z = 0$ (the Supergalactic plane, middle row), and $Z = -20 \, h^{-1}$Mpc (bottom row). Contour spacing is 0.2, the heavier contour is $\delta = 0$, solid contours mark $\delta > 0$ and dashed contours $\delta < 0$. The density is also indicated by shading. Also drawn is the difference field in units of the error, where the contour spacing is unity (right-hand column). The maps are drawn out to a radius of $80 \, h^{-1}$Mpc, and the very thick contour marks the boundaries of the $R_e = 40 \, h^{-1}$Mpc comparison volume.

column is the difference between the two, $\Delta\delta \equiv \delta_P - (b_I^{-1}\delta_I + c)$, divided by the effective error

$$\sigma_e \equiv (\sigma_P{}^2 + b_I^{-2}\sigma_I{}^2)^{1/2} \tag{35}$$

where b_I and c are the best-fit values for the standard volume (§ 6.3, Table 1). The contouring is described in the figure caption.

Although the *IRAS* map is noticeably less noisy than the POTENT map, they both reveal the same main features. The prominent density peak near $(X, Y, Z) = (-30, 15, 0)$ h^{-1}Mpc in the *IRAS* field is centered on the Hydra-Centaurus supercluster, which we associate with the Great Attractor. It is elongated along the Y direction because, under G12 smoothing, the Virgo cluster [located roughly at $(15, 0, 0)$] is an extension of the GA. The GA is clearly seen in the POTENT field roughly in the same region of space, except that it peaks at a different location $(-45, 0, 0)$ than the *IRAS* peak. The GA is also clearly visible in the $Z = -20\,h^{-1}$Mpc plane, where the peak appears in a similar location $(-30, 15)$ in both panels. There is an apparently statistically significant difference between the two, centered at $(-45, 0, 0)$, but this lies in the Zone of Avoidance, where the sampling is very sparse and the sampling gradient bias has the strongest effect. This is why the comparison volume does not extend to this region. Note in particular that very little of the volume is discrepant by more than $2\,\sigma$ (cf., Figure 8 below). In the $Z = +20\,h^{-1}$Mpc plane, the GA lies outside the comparison volume.

The extended density enhancement on the right of the $Z = 0$ maps is the Perseus-Pisces supercluster, which has two peaks, both outside the comparison volume in either map. The PP and GA superclusters are separated by the large Sculptor void, which is similar in the two maps, both at $Z = 0$ and $Z = -20\,h^{-1}$Mpc. Above the Supergalactic plane ($Z = +20\,h^{-1}$Mpc), the density field is dominated by a large void, prominent in both POTENT and *IRAS*. Another density peak, far outside our comparison volume, is the Coma supercluster, centered at $(5, 75, 0)$ in the *IRAS* map, and at $(20, 50, 0)$ in the POTENT reconstruction. This difference need not worry us as the POTENT reconstruction is known to be extremely noisy there.

6.2. Goodness of Fit

Only after we convince ourselves that the data are consistent with the assumed model of GI and linear biasing can we estimate the parameter b_I^{-1} and interpret it properly. We therefore make an effort to assess goodness of fit by more quantitative investigations of the residuals $\Delta\delta$ between the POTENT and *IRAS* fields. We do it in three ways; by checking the Gaussianity of the residuals, by evaluating the mean amplitude of the residuals, and by searching for possible spatial correlations between them.

A visual impression can be gained from Figure 7, which shows a scatter plot of δ_P versus δ_I at the grid points within the standard volume. Recall that the ~ 2000 points are highly correlated because of the G12 smoothing, with only $N_{\rm eff} \sim 18$ independent sub-volumes.
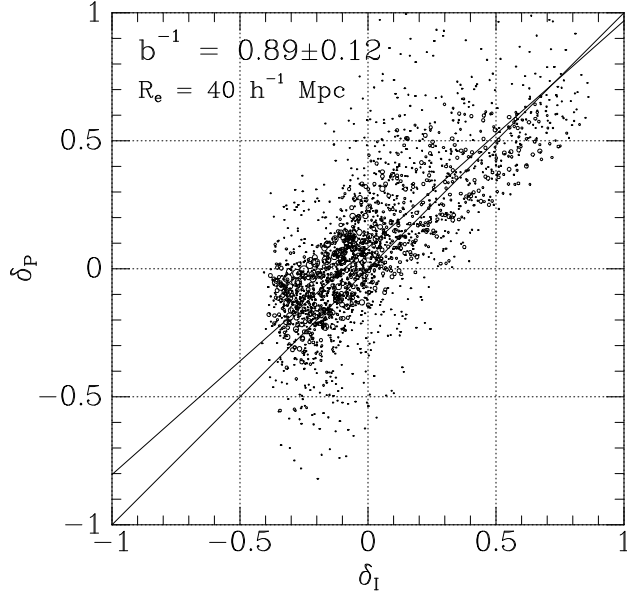
Fig. 7.— POTENT versus *IRAS* G12 density fields from the real data, at grid points within the standard comparison volume. The symbol size is inversely proportional to the effective error. The solid line is the best-fit minimizing equation (33).

This gives rise to coherent clouds of points in the scatter diagram. Notice the qualitative similarity between this figure and the equivalent one for a single, arbitrary, pair of mock realizations, Figure 3.

If we have estimated our errors accurately, and if they are Gaussian distributed, then we expect the scatter $\Delta\delta/\sigma_e$ to be Gaussian distributed with a mean of zero and a variance of unity. The observed differential and cumulative distribution function of this quantity are shown in Figure 8 in comparison with the expected Gaussian. The Gaussian is a reasonable match to the observed distribution; the observed standard deviation of 1.03 (*i.e.*, the square root of the reduced χ^2 statistic S, Table 1, right section) is only slightly different from unity. Note in particular the reasonable fit in the tails of the distribution. Our conclusion is that the distribution of residuals is fairly consistent with a Gaussian.

Figure 9 shows the observed value of S within the standard volume in comparison with the distribution of S over the 90 pairs of mock realizations of the data, representing the null hypothesis of GI and linear biasing. Note that this distribution has no extended tails, and would be well-fit by a Gaussian. The observed value of S falls 1σ from the center of the distribution of S in the mocks (Table 1), which is a confirmation for the goodness of fit. For the smaller and larger comparison volumes, the observed value of S is at 1.6σ and 1.5σ respectively, indicating a weaker, but still acceptable goodness of fit.

It is also important to rule out possible spatial correlations between the residuals, which could be indicative of large-scale systematic errors in the POTENT or *IRAS* density fields.
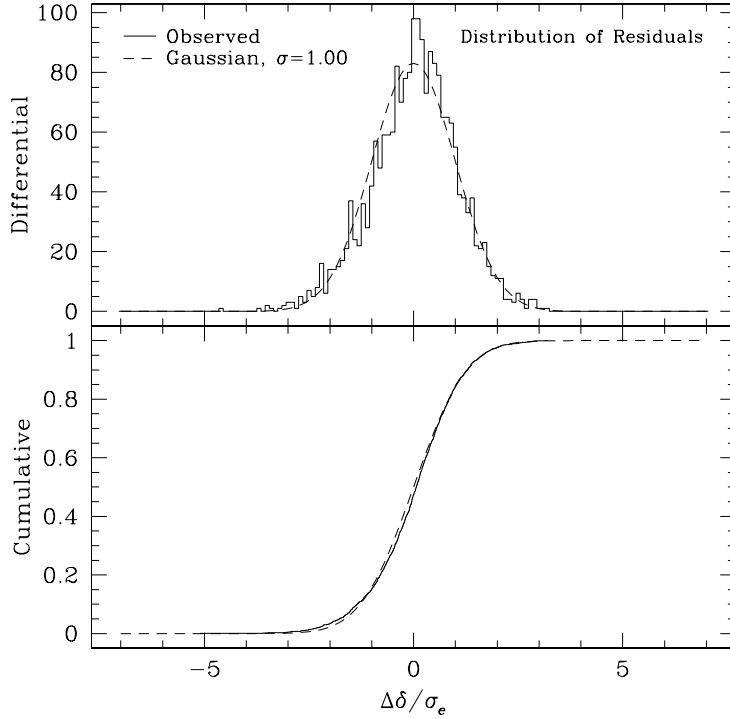
Fig. 8.— The probability distribution function of the normalized residuals, $\Delta\delta/\sigma_e$ (solid), in comparison with the expected Gaussian of $\sigma = 1.00$ (long-dash). Top: differential. Bottom: cumulative.

We define a correlation function of residuals by

$$S(r) = \left\langle \frac{\Delta\delta(\boldsymbol{x})\,\Delta\delta(\boldsymbol{x}+\boldsymbol{r})}{\sigma_e(\boldsymbol{x})\,\sigma_e(\boldsymbol{x}+\boldsymbol{r})} \right\rangle_{\boldsymbol{x}} \quad , \tag{36}$$

where the average is over pairs of grid points in the comparison volume, with pair separation in the range $(r, r+\Delta r)$, for $\Delta r = 5\,h^{-1}\mathrm{Mpc}$. Figure 10 shows this statistic for the real data in comparison with the 90 pairs of mock catalogs. The correlation function at zero separation coincides with the quantity S of equation (34) by definition, and it is a factor of two smaller at $r \sim 12\,h^{-1}\mathrm{Mpc}$, roughly coinciding with the smoothing length. No statistically significant deviation is seen between the residual correlations of the real data and the simulations, at any separation. Thus, there is no evidence for residual correlations in the real data beyond those induced by the finite smoothing.

Finally, the values for the offset between the fields are found to be $c = 0.08 \pm 0.07$, 0.08 ± 0.06 and 0.09 ± 0.07 for the three comparison volumes respectively, where the errors are the standard deviations from the mock catalogs. A priori, one should not be surprised to find a small offset between the POTENT and *IRAS* density fields, because their mean densities were determined in somewhat different volumes and in different ways. These offsets
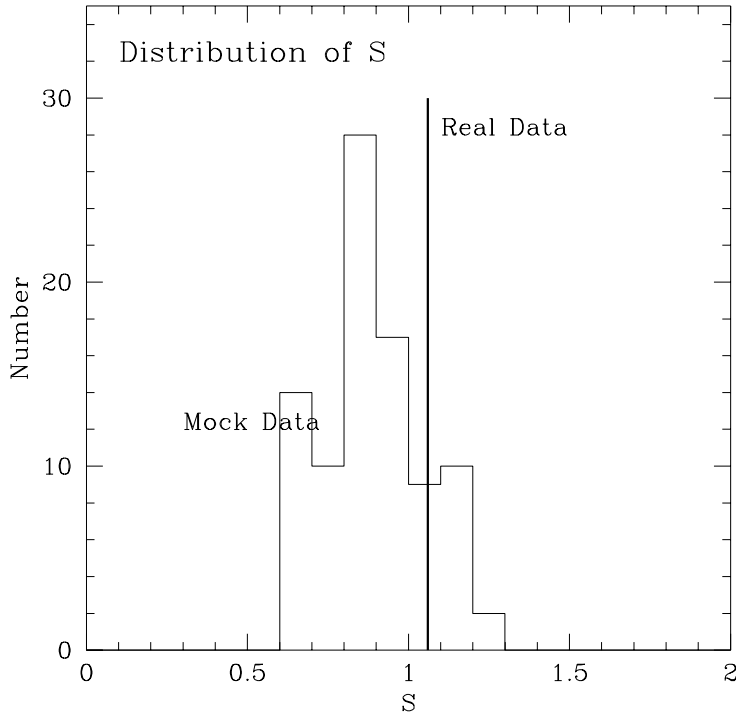
Fig. 9.— The observed value of S within the standard volume (thick line) in comparison with the model distribution of S over the pairs of mock realizations of the data (histogram), within the same volume.

are consistent with the mean and scatter found in the mock catalogs (§ 5.2), indicating that the small offsets are not significant.

Based on the above tests, we conclude that the *IRAS* and POTENT density field data, at G12 smoothing and within the comparison volumes, are *consistent* with the hypotheses of a GI relation between the density and velocity fluctuation fields, and a linear biasing relation between the density fluctuation fields of galaxies and mass.

6.3.   The Value of β_I

Having established a reasonable goodness-of-fit, we are now in a position to derive the value of b_I^{-1}. As we have assumed $\Omega = 1$ in the determination of δ_P, this immediately gives us β_I. The results are given in Table 1. In the standard volume we obtain $\beta_I = 0.89 \pm 0.12$.

Had we been in the linear regime and using purely linear analysis, the simple relation $\mathbf{\nabla} \cdot \mathbf{v} = -\beta_I \delta_I$ would have been strictly valid, and the slope of the best-fit line quoted would have been a direct measure of β_I. Since we have included mildly nonlinear corrections both in
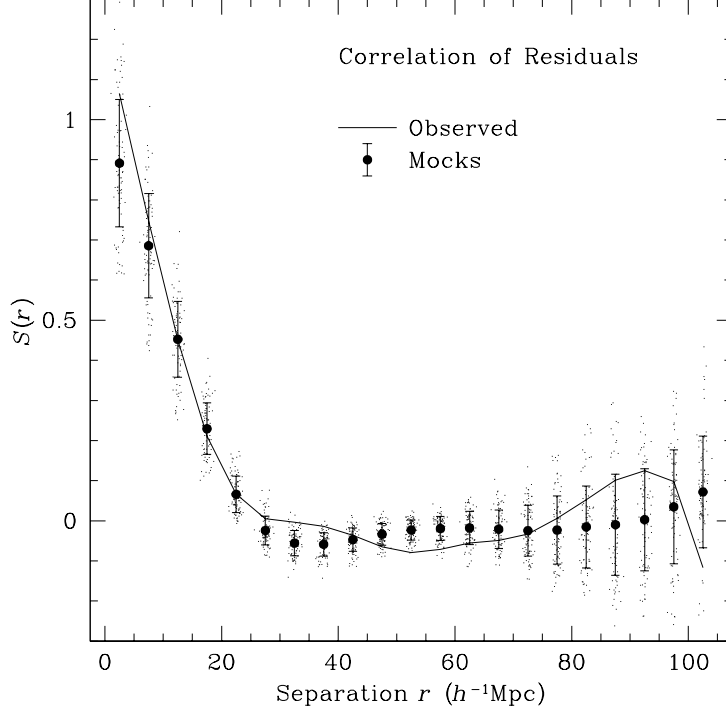
Fig. 10.— The correlation function $S(r)$ for the real data (solid), and for the 90 pairs of mock catalogs (dots). The large symbols with error bars are the mean and standard deviation of the results from the mock catalogs at each bin of separation.

the POTENT and the *IRAS* analyses, we may expect the actual relation between δ_P and δ_I to deviate slightly from the linear relation. What we have done, in fact, was to assume $\Omega = 1$ in the POTENT reconstruction (and, less importantly, $\Omega = b_I = 1$ in the *IRAS* conversion from redshift to real space), and then determine b_I [and thus $\beta_I = f(\Omega)/b_I$] from the best fit line of δ_P versus δ_I. The fact that we ended up with a value of β_I that is consistent at the $\sim 1\sigma$ level with the assumed value of unity indicates that this is a self-consistent solution for Ω and b_I near unity.

In order to validate our result for the more general case of $\beta_I \sim 1$ where Ω and b_I may differ from unity, one should redo the N-body simulation of the mock catalogs for values of $\Omega \neq 1$. In the present paper, we estimate the errors for $\Omega < 1$ by scaling the errors determined from the $\Omega = 1$ simulation. We first redo the comparison analysis in the standard volume with assumed values of $\Omega = 0.3$ and $b_I = 0.5$, which also correspond to $\beta_I = 1$. This makes a difference mostly in the POTENT reconstruction. We crudely scale the errors σ_P in proportion to $f(\Omega)^{-1}$. For this case we obtain $\beta_I = 1.00 \pm 0.18$ from our fit to the real data in our standard volume. The quoted error is now only a rough estimate based on the increase in σ_e. This result is consistent with the value $\beta_I = 0.89 \pm 0.12$, obtained in our standard case of $\Omega = b_I = 1$. This confirms our assessment that we are indeed measuring β_I, and are not very sensitive to the actual values of Ω and b_I that give rise to that value of β_I.

Finally, we try two cases where the assumed β_I is ~ 0.5 rather than unity. When we start with $\Omega = 1$ and $b_I = 2$, we obtain $\beta_I = 0.93 \pm 0.12$. When we start with $\Omega = 0.3$ and $b_I = 1$, we obtain $\beta_I = 0.90 \pm 0.18$ where the error is roughly scaled as before. In both cases, the resultant β_I are significantly larger than the "assumed" value of $\beta_I = 0.5$. This test demonstrates that the result is not determined by the assumed input and the method is discriminatory; we can rule out low values of $\beta_I \leq 0.5$. Note, however, that the rejection of the case $b_I = 2$ is stronger than the rejection of the case $\Omega = 0.3$, because the POTENT errors become larger when Ω is lower.

The above tests strengthen our confidence in the generality of our result from the standard case, of $\beta_I \simeq 0.89$. However, the error estimate of ± 0.12 is strictly valid only near $\Omega \sim 1$. If Ω is lower, the error is expected to be larger. The calibration of the method and the error estimate can be fine-tuned further using full POTENT and *IRAS* reconstructions from mock catalogs drawn from simulations with different values of Ω and b_I, which is work in progress. However, since we have no reason to suspect any significant change in the result, we defer this further testing to a subsequent paper.

7.   DISCUSSION AND CONCLUSIONS

We have compared the density of mass and light as reconstructed from the Mark III catalog of peculiar velocities and the *IRAS* 1.2 Jy redshift survey, with Gaussian smoothing of $12\,h^{-1}$Mpc and within volumes of effective radii $31-46\,h^{-1}$Mpc. Our two main conclusions are:

- The data are consistent with gravitational instability theory in the mildly nonlinear regime and a linear biasing relation between galaxies and mass on these large scales.

- The value of the corresponding β parameter at this smoothing scale is $\beta_I = 0.89 \pm 0.12$. The relative robustness of this value to changes in the comparison volume suggests that the cosmic scatter is no greater than ± 0.1.

This result is consistent with the simplest cosmological model of an Einstein – de Sitter universe ($\Omega = 1$, $\Lambda = 0$) with unbiased *IRAS* galaxies, and it also permits somewhat lower values of Ω and b_I. However, values as low as $\Omega \sim 0.3$ would require large-scale anti-biasing of $b_I < 0.75$ (with 95% confidence) — a phenomenon that is not easily reproduced in theoretical simulations (e.g., Cen & Ostriker 1992; Evrard, Summers, & Davis 1994).

Although our analysis uses an extension of the Zel'dovich approximation, we have not tried to directly measure nonlinear effects in these data to break the degeneracy between Ω and b_I: the effects are weak (PI93), and are of the same order as possible nonlinearities in the biasing relation. An interesting area for future work is to see if we can put an *upper* limit on the degree of nonlinearity; if we can, this puts a lower limit on b_I, and therefore a lower limit on Ω.

As explained in § 1, the current analysis is a significantly improved version of the PI93 analysis done with the earlier Mark II and *IRAS* 1.936 Jy data, and of the Hudson *et al.* (1995) comparison to optical galaxies. The result of PI93 was $\beta_I = 1.28^{+0.75}_{-0.59}$ at 95% confidence. The new result is lower, but only by about 1.3σ. The current analysis is superior in many respects. The improvements in the data include denser sampling of a larger volume, a careful procedure of selection, calibration and merging of the TF data sets, and a better correction for Malmquist bias. The methods of reconstruction were considerably improved, using new techniques and with extensive testing and error analysis using realistic mock catalogs drawn from simulations. The optimization of the comparison method using the mock catalogs led to an unbiased estimate of β with well-defined statistical errors.

The quantity β has been measured using a variety of techniques from observational data sets (as discussed in the references in § 1); we here discuss other estimates of β_I from the Mark III and *IRAS* 1.2 Jy data.

Estimates of β_I from *redshift distortions* in redshift surveys are almost all within 2σ of our current result (Peacock & Dodds 1994, $\beta_I = 1.0 \pm 0.2$; Fisher *et al.* 1994b, $\beta_I = 0.45^{+0.3}_{-0.2}$; Fisher, Sharf, & Lahav 1994c, $\beta_I = 1.0 \pm 0.3$; Cole, Fisher, & Weinberg 1995, $\beta_I = 0.5 \pm 0.15$; Hamilton 1995, $\beta_I = 0.7 \pm 0.2$; Heavens & Taylor 1995, $\beta_I = 1.1 \pm 0.3$; Fisher & Nusser 1996, $\beta_I = 0.6 \pm 0.2$). The scatter between these results is mostly due to differences in method of analysis, but cosmic scatter, nonlinear effects, and complications in the biasing scheme probably also play a role.

Our current analysis is a comparison at the density level (a $\delta - \delta$ comparison). Alternative analyses that performed the comparison at the *velocity* level ($v - v$ comparisons) typically yielded somewhat lower values for β_I, ranging from 0.5 to 0.9: (Strauss 1989, $\beta_I \simeq 0.8$; Kaiser *et al.* 1991, $\beta_I = 0.86^{+0.2}_{-0.15}$; Roth 1994, $\beta_I = 0.6 \pm 0.2$; Nusser & Davis 1994, $\beta_I = 0.6 \pm 0.2$; Davis *et al.* 1997, $\beta_I = 0.6 \pm 0.2$; Willick *et al.* 1997b, $\beta_I = 0.49 \pm 0.07$). We focus here on the two most recent and sophisticated analyses of the $v - v$ type, termed ITF and VELMOD.

The ITF analysis (Davis *et al.* 1997) is a mode-by-mode comparison of velocity-field models, based on expansion in spherical harmonics and radial Bessel functions, that were fitted in redshift space to a processed version of the inverse TF data from the Mark III catalog and the *IRAS* 1.2 Jy redshift survey. The ITF analysis finds a poor fit between the velocity fields and the model, in the form of a distance-dependent dipole at large distances, in apparent contrast with the good fit obtained here. One possible explanation for this difference is an error in one of the large-scale velocity fields, or both, that does not propagate to their local derivatives, the density fields. A differential offset in the zero point of the TF relation between data sets of the Mark III catalog, which generates an artificial coherent bulk motion, is one such possibility. An inaccurate definition of the Local Group frame in the *IRAS* reconstruction, or missing density data beyond the sampled volume, are other possible sources of error in the velocities. Another source of error is the somewhat arbitrary adjustments that had to be made to the Mark III data sets when combining them into one system for the ITF analysis (cf., the discussion in Davis *et al.* 1997). If one ignores the poor

fit, the ITF analysis yields the formal estimate $\beta_I = 0.6 \pm 0.2$, $\sim 1\sigma$ lower than our current result. The local nature of the density-density comparison, and the resulting better goodness of fit, argue that it may provide a more reliable estimate of β_I.

The VELMOD analysis (Willick *et al.* 1997b) is a high-resolution $v - v$ comparison at a small smoothing scale of $3\,h^{-1}$Mpc, using about one quarter of the Mark III data within a volume of radius $< 30\,h^{-1}$Mpc, and applying a sophisticated likelihood analysis in the linear regime. The comparison revealed a good fit and an estimate of $\beta_I = 0.49 \pm 0.07$. We can see several possible reasons for the difference in the estimates of β_I.

It is possible that the difference originates from systematic errors that were somehow overlooked despite the successful testing using the same mock catalogs. For example, the VELMOD analysis assumes pure linear theory, and may thus suffer from nonlinear effects that were not taken into account. In fact, it was never fully understood how the linear analysis of VELMOD did well at G3 smoothing where nonlinear effects are expected to be important. One suspicion is that the mock simulation is too smooth on small scales.

Another source for the difference may be the partial data used in the VELMOD comparison, where the comparison is in fact dominated by data within $\sim 20\,h^{-1}$Mpc. The difference in β_I can thus partly reflect cosmic scatter in b_I or in the local effective Ω. Table 1 shows that β_I increases by an insignificant 1σ from the smallest to the largest comparison volume; the VELMOD analysis also found no statistically significant growth of β_I with scale. This allows us to put an upper limit on the cosmic scatter of order 0.1.

Finally, a possible explanation for the difference in β_I is *scale dependence* of the biasing relation between Gaussian smoothings of 3 and $12\,h^{-1}$Mpc, which would be associated with non-linear biasing. Some support for such a trend is found by the SIMPOT analysis (Nusser & Dekel 1997), which fits a parametric model of velocity and density fields and β simultaneously to the peculiar velocity and redshift data (a $v - \delta$ comparison). This analysis yields $\beta_I = 1.0 \pm 0.15$ for G12 smoothing, and lower values of β_I for smaller smoothing scales, in qualitative agreement with the comparison of the results of the current paper and of VELMOD. Current theoretical simulations indicate possible scale dependence in the biasing relation between scales of one to a few megaparsecs (*e.g.*, Kauffman *et al.* 1997; cf., Mo & White 1996), but it remains to be seen whether such a trend can arise on scales of 6 to $12\,h^{-1}$Mpc. The biasing scheme, which could be nontrivial in several ways (see Dekel & Lahav 1997), is clearly a bottle-neck in the effort to measure the cosmological parameter Ω via β — the biasing is inevitably involved whenever galaxy-density data are used.

The largest source of error in this analysis lies in the peculiar velocity data, and thus improved peculiar velocity datasets, with better control of systematic errors, denser sampling, and more complete sky coverage, will be of great importance for this work (cf., the reviews of Strauss 1997a; Giovanelli 1997). It will be interesting to make the comparison of the POTENT maps with the optical redshift data of Santiago *et al.* (1995, 1996), although a proper calculation of the errors in the optical density field will be somewhat more difficult. Finally, more work lies ahead in testing the robustness of our results to the assumed value

of Ω (see the discussion in § 6.3), and carrying out the analysis at other smoothing lengths to look for nonlinear effects.

REFERENCES

Aaronson, M., *et al.* 1982, ApJS, 50, 241

Bardeen, J., Bond, J. R., Kaiser, N., & Szalay, A. 1986, ApJ, 304, 15

Bertschinger, E., & Dekel, A. 1989, ApJ, 336, L5

Bertschinger, E., & Gelb, G. 1991, Computers in Physics, 5, 164

Burstein, D. 1989, privately circulated computer files (Mark II)

Cen, R.Y., & Ostriker, J.P. 1992, ApJ, 399, L113

Cole, S., Fisher, K. B., & Weinberg, D. 1995, MNRAS, 275, 515

Courteau, S. 1992, PhD. Thesis, University of California, Santa Cruz

Davis, M., Nusser, A., & Willick, J. A. 1996, ApJ, 473, 22

Dekel, A. 1994, ARA&A, 32, 371

Dekel, A. 1997a, in *Galaxy Scaling Relations: Origins, Evolution and Applications* (ESO workshop, November 1996), ed. L. da Costa (Springer) in press (astro-ph/9705033)

Dekel, A. 1997b, in *Formation of Structure in the Universe*, eds. A. Dekel & J.P. Ostriker (Cambridge Univ. Press) in press

Dekel, A., Bertschinger, E., & Faber, S. M. 1990, ApJ, 364, 349 (DBF)

Dekel, A., Bertschinger, E., Yahil, A., Strauss, M. A., Davis, M., & Huchra J. P. 1993, ApJ, 412, 1 (PI93)

Dekel, A., Burstein, D., & White, S.D.M. 1997, in *Critical Dialogues in Cosmology*, ed. N. Turok (Singapore: World Scientific), in press (astro-ph/9611108)

Dekel, A., Eldar, A., Kolatt, T., Willick, J. A., Faber, S. M., Courteau, S., & Burstein, D. 1997, in preparation

Dekel, A., & Lahav, O. 1997, in preparation

Dekel, A., & Rees, M. J. 1987, Nature, 326, 455

Dressler, A. 1980, ApJ, 236, 351

Eldar, E., Dekel, A., & Willick, J.A. 1997, in preparation

Evrard, A.E., Summers, F.J., & Davis, M. 1994, ApJ, 422, 11

Fisher, K. B., Davis, M., Strauss, M. A., Yahil, A., & Huchra J. P. 1993, ApJ, 402, 42

Fisher, K. B., Davis, M., Strauss, M. A., Yahil, A., & Huchra, J. P. 1994a, MNRAS, 266, 50

Fisher, K. B., Davis, M., Strauss, M. A., Yahil, A., & Huchra, J. P. 1994b, MNRAS, 267, 927

Fisher, K. B., Huchra J. P., Strauss, M. A., Davis, M., Yahil, A., & Schlegel D. 1995, ApJS, 100, 69

Fisher, K. B., & Nusser, A. 1996, MNRAS, 279, L1

Fisher, K. B., Scharf, C. A., & Lahav, O. 1994c, MNRAS, 266, 219

Ganon, G., Dekel, A., Mancinelli, P., & Yahil, A. 1997, in preparation

Giovanelli, R. 1997, in *Critical Dialogues in Cosmology*, ed. N. Turok (Singapore: World Scientific), in press (astro-ph/9610129)

Hamilton, A. J. S. 1995, in *Clustering in the Universe*, Proc. 30$^{\text{th}}$ Rencontres de Moriond, ed. S. Maurogordato, C. Balkowski, C. Tao, and J. Trân Thanh Vân (Gif-sur-Yvette: Editions Frontières), 143

Han, M.-S., & Mould, J. R. 1992, ApJ, 396, 453

Heavens, A. F., & Taylor, A. N. 1995, MNRAS, 275, 483

Hoffman, Y., & Ribak, E. 1991, ApJ, 380, L5

Hudson, M. J. 1993, MNRAS, 265, 43

Hudson, M. J. 1994, MNRAS, 266, 475

Hudson, M. J., Dekel, A., Courteau, S., Faber, S. M., & Willick, J. A. 1995, MNRAS, 274, 305

Kaiser, N. 1987, MNRAS, 227, 1

Kaiser, N., Efstathiou, G., Ellis, R., Frenk, C., Lawrence, A., Rowan-Robinson, M., & Saunders, W. 1991, MNRAS, 252, 1

Kauffman, G., Nusser, A., & Steinmetz, M. 1997, MNRAS, 286, 795

Kolatt, T., Dekel, A., Ganon, G., & Willick, J. A. 1996, ApJ, 458, 419

Lahav, O., Fisher, K. B., Hoffman, Y., Scharf, C. A., & Zaroubi, S. 1994, ApJ, 423, L93

Landy, S. D., & Szalay, A. S. 1992, ApJ, 391, 494

Lawley, D. N., & Maxwell, A. E. 1971, *Factor Analysis as a Statistical Method* (New York: Elsevier)

Lynden-Bell, D., Faber, S. M., Burstein, D., Davies, R. L., Dressler, A., Terlevich, R. J., & Wegner, G. 1988, ApJ, 326, 19

Mancinelli, P. J., & Yahil, A. 1995, ApJ, 452, 75

Mancinelli, P. J., Yahil, A., Ganon, G., & Dekel, A. 1994, in *Cosmic Velocity Fields*, eds. F. Bouchet & M. Lachiéze-Rey (Gif-sur-Yvette: Editions Frontières), 215

Mathewson, D. S., Ford, V. L, & Buchhorn, M. 1992, ApJS, 81, 413

Mo, H., & White, S.D.M. 1996, MNRAS, 282, 347

Nusser, A., & Davis, M. 1995, MNRAS, 276, 1391

Nusser, A., & Dekel, A. 1997, in preparation

Nusser, A., Dekel, A., Bertschinger, E., & Blumenthal, G. R. 1991, ApJ, 379, 6

Peacock, J. A., & Dodds, S. J. 1994, MNRAS, 267, 1020

Press, W. H., Flannery, B. P., Teukolsky, S. A., & Vetterling, W. H. 1992, *Numerical Recipes, the Art of Scientific Computing* (Second Edition) (Cambridge: Cambridge University Press)

Roth, J. R. 1994, in *Cosmic Velocity Fields*, eds. F. Bouchet & M. Lachiéze-Rey (Gif-sur-Yvette: Editions Frontières), 233

Santiago, B. X., Strauss, M. A., Lahav, O., Davis, M., Dressler, A., & Huchra, J. P. 1995, ApJ, 446, 457

Santiago, B.X., Strauss, M.A., Lahav, O., Davis, M., Dressler, A., & Huchra, J.P. 1996, ApJ, 461, 38

Schlegel, D. 1995, PhD. Thesis, University of California, Berkeley

Shaya, E. J., Peebles, P. J. E., & Tully, R. B. 1995, ApJ, 454, 15

Strauss, M. A. 1989, Ph.D. Thesis, University of California, Berkeley

Strauss, M. A. 1997a, in *Critical Dialogues in Cosmology*, ed. N. Turok (Singapore: World Scientific), in press (astro-ph/9610032)

Strauss, M. A. 1997b, in *Formation of Structure in the Universe*, eds. A. Dekel & J.P. Ostriker (Cambridge Univ. Press) in press (astro-ph/9610033)

Strauss, M. A., Davis, M., Yahil, A., & Huchra, J. P. 1990, ApJ, 361, 49

Strauss, M. A., Huchra, J. P., Davis, M., Yahil, A., Fisher, K. B., & Tonry, J. 1992b, ApJS, 83, 29

Strauss, M. A., & Willick, J. A. 1995, Phys. Rep., 261, 271

Strauss, M. A., Yahil, A., Davis, M., Huchra J. P., & Fisher, K. B. 1992a, ApJ, 397, 395

Tormen, G., & Burstein, D. 1995, ApJS, 96, 123

Willick, J. 1991, PhD. thesis, University of California, Berkeley

Willick, J. A., Courteau, S., Faber, S.M., Burstein, D., & Dekel, A. 1995, ApJ, 446, 12

Willick, J. A., Courteau, S., Faber, S.M., Burstein, D., Dekel, A., & Kolatt, T. 1996, ApJ, 457, 460

Willick, J. A., Courteau, S., Faber, S.M., Burstein, D., Dekel, A., & Strauss, M. A. 1997a, ApJS, 109, 333

Willick, J. A., Strauss, M. A., Dekel, A., & Kolatt, T. 1997b, ApJ, in press (astro-ph/9612240)

Yahil, A. 1994, in *Evolution of the Universe and its Observational Quest*, ed. K. Sato (Tokyo: Universal Academy Press), 197

Yahil, A., Strauss, M. A., Davis, M., & Huchra J. P. 1991, ApJ, 372, 380

Zaroubi, S., Hoffman, Y., Fisher, K. B., & Lahav, O. 1995, ApJ, 449, 446

Zel'dovich, Y. B. 1970, A&A, 5, 84